

# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 21  
2015  
№ 9

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ  
Издательство "Новые технологии"

## СОДЕРЖАНИЕ

### ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

- Родионов А. Н. Периодические данные: идентификация, моделирование и представление в базах данных . . . . . 643  
Катуева Я. В., Назаров Д. А. Методы параметрического синтеза на основе точного представления области работоспособности . . . . . 651

### МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

- Каширина И. Л., Львович Я. Е., Сорокин С. О. Модели и численные методы оптимизации формирования эффективной сетевой системы с кластерной структурой . . . . . 657  
Бронштейн Е. М., Копылов И. Е. Оптимизационная задача транспортной логистики с простыми маршрутами передвижения . . . . . 662  
Моисеев А. А. Моделирование химической стабильности жидкости методом теории подобия . . . . . 666

### ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

- Асратян Р. Э., Лебедев В. Н., Орлов В. Л. Организация защищенных каналов взаимодействия на основе применения протокола HTTPS в прокси-серверах . . . . . 670  
Саак А. Э. Диспетчеризация массивов заявок суммарной ресурсной меры, равной квадрату целого числа . . . . . 675

### ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ

- Потапов В. П., Попов С. Е. Многопоточный алгоритм конвертации чередования данных каналов гиперспектральных снимков . . . . . 680  
Дворников С. В., Пшеничников А. В., Манаенко С. С. Помехоустойчивая модель сигнала КАМ-16 с трансформированным созвездием . . . . . 685  
Светушков Н. Н. Параллельный метод струн для численного решения нелинейных задач теплопроводности . . . . . 689

### СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

- Соловьев Р. А., Тельпухов Д. В., Рухлов В. С., Поперечный П. С. Особенности проектирования модулярных многоходовых сумматоров с помощью современных САПР . . . . . 694  
Гураий М. М., Русаков С. Г., Алиев Р. Р. Применение методов схемотехнического моделирования для анализа биоэлектрических процессов в многоклеточных системах . . . . . 699  
Ноженкова Л. Ф., Исаева О. С., Грузенко Е. А., Вогоровский Р. В., Колдырев А. Ю., Евсюков А. А. Комплексная поддержка конструирования бортовых систем контроля и управления космических аппаратов на основе интеллектуальной имитационной модели . . . . . 706

### ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

- Яшина А. Г., Прозоров Д. Е. Взвешенная косинусная мера векторной модели информационного поиска речевых документов . . . . . 715

Главный редактор:  
СТЕМПКОВСКИЙ А. Л.,  
акад. РАН, д. т. н., проф.

Зам. главного редактора:  
ИВАННИКОВ А. Д., д. т. н., проф.  
ФИЛИМОНОВ Н. Б., д. т. н., с.н.с.

Редакционный совет:  
БЫЧКОВ И. В., акад. РАН, д. т. н.  
ЖУРАВЛЕВ Ю. И.,  
акад. РАН, д. ф.-м. н., проф.  
КУЛЕШОВ А. П.,  
акад. РАН, д. т. н., проф.  
ПОПКОВ Ю. С.,  
чл.-корр. РАН, д. т. н., проф.  
РУСАКОВ С. Г.,  
чл.-корр. РАН, д. т. н., проф.  
РЯБОВ Г. Г.,  
чл.-корр. РАН, д. т. н., проф.  
СОЙФЕР В. А.,  
чл.-корр. РАН, д. т. н., проф.  
СОКОЛОВ И. А., акад.  
РАН, д. т. н., проф.  
СУЕТИН Н. В., д. ф.-м. н., проф.  
ЧАПЛЫГИН Ю. А.,  
чл.-корр. РАН, д. т. н., проф.  
ШАХНОВ В. А.,  
чл.-корр. РАН, д. т. н., проф.  
ШОКИН Ю. И.,  
акад. РАН, д. т. н., проф.  
ЮСУПОВ Р. М.,  
чл.-корр. РАН, д. т. н., проф.

Редакционная коллегия:  
АВДОШИН С. М., к. т. н., доц.  
АНТОНОВ Б. И.  
БАРСКИЙ А. Б., д. т. н., проф.  
ВАСЕНИН В. А., д. ф.-м. н., проф.  
ВИШНЕКОВ А. В., д. т. н., проф.  
ГАЛУШКИН А. И., д. т. н., проф.  
ДИМИТРИЕНКО Ю. И., д. ф.-м. н., проф.  
ДОМРАЧЕВ В. Г., д. т. н., проф.  
ЗАБОРОВСКИЙ В. С., д. т. н., проф.  
ЗАГИДУЛЛИН Р. Ш., к. т. н., доц.  
ЗАРУБИН В. С., д. т. н., проф.  
КАРПЕНКО А. П., д. ф.-м. н., проф.  
КОЛИН К. К., д. т. н., проф.  
КУЛАГИН В. П., д. т. н., проф.  
КУРЕЙЧИК В. М., д. т. н., проф.  
КУХАРЕНКО Б. Г., к. ф.-м. н., доц.  
ЛЬВОВИЧ Я. Е., д. т. н., проф.  
МИХАЙЛОВ Б. М., д. т. н., проф.  
НЕЧАЕВ В. В., к. т. н., проф.  
ПОЛЕЩУК О. М., д. т. н., проф.  
СОКОЛОВ Б. В., д. т. н., проф.  
ТИМОНИНА Е. Е., д. т. н., проф.  
УСКОВ В. Л., к. т. н. (США)  
ФОМИЧЕВ В. А., д. т. н., проф.  
ЧЕРМОШЕНЦЕВ С. Ф., д. т. н., проф.  
ШИЛОВ В. В., к. т. н., доц.

Редакция:  
БЕЗМЕНОВА М. Ю.  
ГРИГОРИН-РЯБОВА Е. В.  
ЛЫСЕНКО А. В.  
ЧУГУНОВА А. В.

Информация о журнале доступна по сети Internet по адресу <http://novtex.ru/IT>.

Журнал включен в систему Российского индекса научного цитирования.

Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

## CONTENTS

### INTELLIGENT SYSTEMS AND TECHNOLOGIES

- Rodionov A. N.** Periodical Data: Identification, Modeling and Representation in Databases . . . . . 643
- Katueva Ya. V., Nazarov D. A.** The Methods of Parametric Synthesis Based on Grid Representation of a Region of Acceptability . . . . . 651

### MODELING AND OPTIMIZATION

- Kashirina I. L., Lvovich Ya. E., Sorokin S. O.** Models and Numerical Methods Optimization of Formation Effective Network Systems Cluster Structure . . . . . 657
- Bronshtcin E. M., Kopylov I. E.** Vehicle Routing Problem with Simple Movement Route . . . . . 662
- Moiseev A. A.** Simulation Chemical Stability of the Liquid by the Method of Similarity Theory . . . . . 666

### COMPUTING SYSTEMS AND NETWORKS

- Asratian R. E., Lebedev V. N., Orlov V. L.** Creating a Secure Communication Channels Based on the Use of the HTTPS in Proxy Servers . . . . . 670
- Saak A. E.** Scheduling of Sets of Multiprocessor Tasks of the Aggregate Resource Measure Equaling the Square of an Integer . . . . . 675

### DIGITAL PROCESSING OF SIGNALS AND IMAGES

- Potapov V. P., Popov S. E.** Multithreaded Algorithm for Band Interleave Conversion of Hyperspectral Images . . . . . 680
- Dvornikov S. V., Pshenichnicov A. V., Manaenko S. S.** Increased Noise Immunity Signal 16-QAM Constellation with Transformed . . . . . 685
- Svetushkov N. N.** Parallel Strings Method for the Numerical Solution of Nonlinear Heat Conduction Problem . . . . . 689

### CAD-SYSTEMS

- Solovyev R. A., Telpukhov D. V., Rukhlov V. S., Poperechny P. S.** Study of Multi-Input Modular Adders Implemented on FPGA and VLSI by Modern CAD Tools . . . . . 694
- Gourary M. M., Rusakov S. G., Aliev R. R.** Application of Circuit Simulation Methods for the Analysis of Bioelectrical Processes in Multicellular Systems. . . . . 699
- Nozhenkova L. F., Isaeva O. S., Gruzenko E. A., Vogorovskiy R. V., Koldyrev A. Yu., Evsyukov A. A.** Complex Support of Designing Onboard Control and Management Systems of Spacecrafts on the Basis of Intellectual Simulation Model . . . . . 706

### APPLIED INFORMATION TECHNOLOGIES

- Yashina A. G., Prozorov D. E.** Spoken Document Retrieval Vector Model Based on Weighted Cosine Measure . . . . . 715

#### Editor-in-Chief:

Stempkovsky A. L., Member of RAS,  
Dr. Sci. (Tech.), Prof.

#### Deputy Editor-in-Chief:

Ivannikov A. D., Dr. Sci. (Tech.), Prof.  
Filimonov N. B., Dr. Sci. (Tech.), Prof.

#### Chairman:

Bychkov I. V., Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Zhuravljov Yu. I., Member of RAS,  
Dr. Sci. (Phys.-Math.), Prof.  
Kuleshov A. P., Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Popkov Yu. S., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Rusakov S. G., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Ryabov G. G., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Soifer V. A., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Sokolov I. A., Member of RAS,  
Dr. Sci. (Phys.-Math.), Prof.  
Suetin N. V.,  
Dr. Sci. (Phys.-Math.), Prof.  
Chaplygin Yu. A., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Shakhnov V. A., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Shokin Yu. I., Member of RAS,  
Dr. Sci. (Tech.), Prof.  
Yusupov R. M., Corresp. Member of RAS,  
Dr. Sci. (Tech.), Prof.

#### Editorial Board Members:

Avdoshin S. M., Cand. Sci. (Tech.), Ass. Prof.  
Antonov B. I.  
Barsky A. B., Dr. Sci. (Tech.), Prof.  
Vasenin V. A., Dr. Sci. (Phys.-Math.), Prof.  
Vishnekov A. V., Dr. Sci. (Tech.), Prof.  
Galushkin A. I., Dr. Sci. (Tech.), Prof.  
Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.  
Domrachev V. G., Dr. Sci. (Tech.), Prof.  
Zaborovsky V. S., Dr. Sci. (Tech.), Prof.  
Zagidullin R. Sh., Cand. Sci. (Tech.), Ass. Prof.  
Zarubin V. S., Dr. Sci. (Tech.), Prof.  
Karpenko A. P., Dr. Sci. (Phys.-Math.), Prof.  
Kolin K. K., Dr. Sci. (Tech.)  
Kulagin V. P., Dr. Sci. (Tech.), Prof.  
Kureichik V. M., Dr. Sci. (Tech.), Prof.  
Kukhareno B. G., Cand. Sci. (Phys.-Math.)  
Lvovich Ya. E., Dr. Sci. (Tech.), Prof.  
Mikhailov B. M., Dr. Sci. (Tech.), Prof.  
Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.  
Poleschuk O. M., Dr. Sci. (Tech.), Prof.  
Sokolov B. V., Dr. Sci. (Tech.)  
Timonina E. E., Dr. Sci. (Tech.), Prof.  
Uskov V. L. (USA), Dr. Sci. (Tech.)  
Fomichev V. A., Dr. Sci. (Tech.), Prof.  
Chermoshentsev S. F., Dr. Sci. (Tech.), Prof.  
Shilov V. V., Cand. Sci. (Tech.), Ass. Prof.

#### Editors:

Bezmenova M. Yu.  
Grigorin-Ryabova E. V.  
Lysenko A. V.  
Chugunova A. V.

Complete Internet version of the journal at site: <http://novtex.ru/IT>.

According to the decision of the Higher Certifying Commission of the Ministry of Education of Russian Federation, the journal is inscribed in "The List of the Leading Scientific Journals and Editions wherein Main Scientific Results of Theses for Doctor's or Candidate's Degrees Should Be Published"

А. Н. Родионов, д-р техн. наук, вед. науч. сотр. e-mail: ran@newmail.ru,  
Вычислительный центр ДВО РАН, г. Хабаровск

## Периодические данные: идентификация, моделирование и представление в базах данных

*Рассматривается полный спектр вопросов моделирования данных, повторяющихся с определенной периодичностью. В первую очередь решается задача по их идентификации. Строится модель, на основании которой могут быть определены параметры циклических процессов, являющихся источником таких данных. Вводятся и исследуются на предмет сопоставимого качества моделирующие структуры и соответствующие им реляционные отношения. Завершает работу рассмотрение вопросов привязки моделируемых структур к единому модельному времени баз данных и перечисление направлений перспективных исследований в этой области.*

**Ключевые слова:** периодические и непериодические циклические данные, тип, экземпляр и клон сущности, многофункциональные циклические структуры данных, темпоральные базы данных

### Введение

Конструирование структур, в которых будут храниться данные о всевозможных "расписаниях" (графиках, планах и подобных им по своему назначению объектах), — одна из типичных задач даталогического моделирования. По сути, в подобных объектах сосредоточена информация о прогнозируемом поведении сущностей. В этих же объектах достаточно часто встречаются как данные, повторяющиеся с некоторой заданной периодичностью (далее "периодические циклические" данные — *pd*), так и данные с произвольным периодом повторения ("непериодические циклические" данные — *nd*).

О наличии или отсутствии *pd* могут косвенно свидетельствовать и ограничения, действующие в тех или иных предметных областях. Вот только два примера подобных ограничений.

*Ограничение 1.* "Лекции по определенной дисциплине читаются один раз в неделю с продолжительностью 2 академических часа".

*Ограничение 2.* "Рейс номер такой-то выполняется 2 раза в неделю по понедельникам и пятницам".

*Лекции и рейсы* здесь претенденты на *pd*, поскольку имеет место повторение и первых, и вторых через равные промежутки времени.

Цикличность данных (если она установлена), период повторения, а также факт идентификации повторяющихся данных дают возможность компактно (используя минимальное число отношений и короткей) представить "расписания" в базах данных.

Разработка схем отношений для простейших "расписаний", когда известны все перечисленные вы-

ше характеристики, не составляет большого труда. В действительности же зачастую только выделить "циклические данные" из массивов документов (которые являются исходным материалом для проектирования) не всегда удается "с первого раза". Например, *лекции* и *рейсы* были отнесены к разряду *pd* скорее интуитивно, чем формально, потому что в обоих случаях негласно присутствуют связанные с ними агенты — дисциплины и воздушные суда соответственно.

Не все очевидно и в отношении периода повторения, например *рейсов*, так как последние выполняются (если судить по ограничению) не только 2 раза в неделю, но и в определенные дни недели. Очевидно, что в последнем случае фиксация в базе данных подобного рода сведений обязательна и "недельный" период требует отдельной, как одно из решений, двукратной привязки к порядковому номеру дня.

Трудности моделирования "циклических" данных обусловлены не только сложностями их обнаружения, формализации и идентификации, но и необходимостью учета времени как фактора их существования. Научное сообщество последнее десятилетие активно развивает темпоральные аспекты концептуального моделирования данных. Они охватывают широкий спектр направлений, включая вопросы отражения траекторий и семантики "движущихся" объектов [1], разработки специфических приемов представления периодичности в виде повторяющихся квантов времени [2], построения и определения областей применения специальных моделей

времени [3, 4], к которым "привязываются" объекты моделей данных. Просматривается тенденция на стандартизацию и ограничение числа базовых временных категорий с унификацией алгоритмов их сопряжения с данными [5]. В то же время аспекты моделирования повторяющихся (циклических) данных, отражающих прогнозные траектории, практически выпадают из поля зрения исследователей.

В работе рассматривается весь комплекс вопросов моделирования, начиная с идентификации собственно циклических данных и заканчивая конструированием и использованием структур, предназначенных для размещения *pd* в базах данных реальных систем.

### "Циклические" данные.

#### Определение и вопросы моделирования

Конечная цель любого моделирования данных — получение качественных концептуальных схем. Качество — комплексная категория и применительно к данным включает множество составляющих, среди которых функциональность (полнота), неизбыточность и отсутствие аномалий [6, 7], равно как и гибкость, простота в использовании, эффективность [8], а также компактность [9], считаются наиболее важными.

При рассмотрении и анализе "циклических" структур ограничимся единственной качественной характеристикой — компактностью, полагая, что соответствие первым трем по умолчанию всегда должно быть достигнуто, а гибкость, простота в использовании и эффективность требуют расширения рамок моделирования, поскольку на их значении оказывают влияние и другие части информационной системы.

Компактными будем полагать те структуры (если более точно, то соответствующие им отношения), которые, если сравнивать их с альтернативными конструкциями, характеризуются меньшей мощностью и разреженностью. (В дальнейшем, при переходах с концептуального уровня моделирования на логический будем ориентироваться на реляционные модели данных, используя соответствующие терминологию и структурную организацию последних.)

Определившись с основным "качественным" критерием, исследуем природу и источники периодических данных в моделируемых предметных областях (универсумах).

На интуитивном уровне "циклические" данные всегда ассоциируются с данными, повторяющимися с некоторой регулярностью. Для их обнаружения в конкретных универсумах сказанного недостаточно. Нужно иметь более точное, формальное определение. Для его получения будем придерживаться основополагающих моментов, касающихся как наиболее общего взгляда на базу данных и ее организацию, так и на устройство организационных систем,

в которых *pd* можно обнаружить наиболее часто. (Ссылка на организационные системы сделана из тех соображений, что при планировании какой-либо деятельности специалисты стремятся к тому, чтобы она протекала по возможности ритмично.)

Поскольку любая предметная область с точки зрения моделирования данных есть не что иное, как набор взаимодействующих между собой сущностей, включая сущности, поступающие извне, сущности, потребляемые и производимые универсумом, а также сущности, покидающие его границы, то база данных, будучи нацеленной на отражение перечисленных аспектов, должна содержать статические и динамические данные, ассоциированные исключительно с сущностями. Фиксируемые в базе данных сведения (факты) могут касаться как текущих моментов времени  $t_c$ , так и будущих периодов. Ретроспективные данные вносятся в базу данных, как правило, при ее создании и играют роль своеобразных начальных условий, отражающих состояние универсума в определенный момент времени  $t_s \leq t_c$ . Текущие данные становятся ретроспективными при переходе к новому транзакционному периоду. Статические базы данных именуются традиционными, а динамические, включающие в том числе обязательные статические аспекты, — темпоральными. Для отражения динамики сущностей ("изменений" сущностей) в базе данных необходимо сосредоточить все факты, касающиеся сведений о местоположениях, состояниях и ролях сущностей.

Из вышеизложенного следует, во-первых, что любое "циклическое данное" должно всегда ассоциироваться с сущностью (принадлежать ей), поскольку ничего, кроме сущностей и характеристик последних, база данных по определению содержать не может. Во-вторых, сама ссылка на "циклическость" указывает на то, что такие данные могут появляться только в темпоральных базах данных, и, следовательно, изменения следует рассматривать в контексте "процессов", что в реальности и происходит. Будем поэтому далее увязывать "циклические" данные с парой <тип сущности, тип процесса> ( $\langle e, p \rangle$ ). Выбор подобного составного объекта для идентификации *pd* предопределяет все последующие теоретические и практические построения.

Кроме традиционных понятий, используемых в даталогическом моделировании, таких как тип и экземпляр [10], позволяющих различать именованные множества и их отдельные экземпляры, для целей настоящего исследования несколько расширим терминологический базис, добавив в него новую категорию "клон" как совокупность идентифицируемых идентичных экземпляров сущностей. С помощью "клонов" будем отличать отдельные, одинаковые экземпляры в последовательностях, которые эти экземпляры будут образовывать.

Если, например, в предметной области "Авиационное сообщение" "рейсы" — это тип, "рейс

№ 8457" — экземпляр, принадлежащий типу "рейс", то "рейс № 8457", выполняемый по вторникам и четвергам", — два клона "рейса № 8457".

Сказанное позволяет в первом приближении определиться со схемой (заголовком) отношения, в котором будут размещаться  $pd$ :

$$R(\text{Цикле}_{ij}) = \{\bar{E}_i^*, \bar{P}_j^*, \bar{T}^*, \bar{r}\}, \quad (1)$$

где  $i$  — индекс типа сущности;  $j$  — индекс типа процесса;  $\bar{E}_i^*$  — атрибут, идентифицирующий экземпляр сущности  $i$ -го типа;  $\bar{P}_j^*$  — атрибут, идентифицирующий процесс  $j$ -го типа;  $\bar{T}^*$  — ссылка на временной период;  $\bar{r}$  — атрибут, "указывающий" на число повторений; надстрочный индекс \* — признак ключевого атрибута. Элементы последнего столбца в отношении, которое соответствует схеме (1), будут содержать непосредственно  $pd$ .

Обозначим круг задач, возникающих при конкретизации и расширении полученного отношения, пригодного для моделирования  $pd$  уже реальных систем.

Обратим внимание на то, что конечной целью является получение моделирующих структур, наилучшим образом, с позиции качества, представляющих  $pd$  в базах данных. В связи с этим первоочередная задача — поиск, идентификация и нахождение параметров  $pd$ , актуальных для моделируемого универсума. Для ее решения потребуется построить формальную модель циклических процессов и разработать на ее основе соответствующий алгоритм.

Следующий шаг — подбор эффективных моделирующих структур, позволяющих фиксировать всевозможные параметры периодичности, и оценка качества таких структур.

Решение проблемы сочетания повторяющихся периодических и непериодических данных, что часто встречается на практике, составляет содержание еще одной задачи моделирования  $pd$ .

Общесистемные аспекты охватывают обязательные вопросы привязки циклических траекторий к осям реального и модельного времени, а также решения по организации хранилищ ретроспективных данных.

В представленной последовательности будет строиться все дальнейшее изложение.

### Модели циклических процессов

Детализации схемы отношения (1) должны предшествовать рассмотрение и формализация по возможности всех наблюдаемых вариантов периодичности, которые встречаются на практике.

Периодичность в информатике, если подходить к этому вопросу формально, будет означать то же самое, что и в математике. Должно существовать число  $m \neq 0$  такое, чтобы выполнялось условие  $f(x + m) = f(x) = y$ . Поскольку  $pd$  ассоциируются с парой  $\langle e, p \rangle$ , то повторяться с известной регулярностью при фиксированном  $e$  должны про-

цессы. Следовательно, циклическую функцию применительно к рассматриваемой задаче можно записать в виде

$$f(t_{pj} + T_{pj}) = f(t) = p_{pj},$$

где  $t_{pj}$  — текущий момент времени начала экземпляра процесса  $p_{pj}$  (подстрочный индекс  $p$  идентифицирует экземпляр процесса  $j$ -го типа);  $T_{pj}$  — продолжительность периода повторения  $p_{pj}$ .

Таким образом, моделирование, по сути, сводится к нахождению  $t_{pj}$  и  $T_{pj}$  для всех экземпляров процессов  $p_{pj}$ , принадлежащих некоторому типу  $P_j$ . (Обозначение  $P_j$  введено, чтобы различать атрибут  $\bar{P}_j$  в отношении (1) и множество процессов  $P_j$ . Далее, моделируя циклические экземпляры  $p_{pj}$ , индекс  $j$  будем опускать, полагая, что все  $p_p$  принадлежат конкретному  $P_j$ .)

Отметим, что каждое множество  $P_j$  характеризуется набором уникальных параметров, часть из которых идентифицирует экземпляры  $p$ , а часть — свойства процесса. Отсюда, отдельные  $p_p$  различаются значениями, принимаемыми параметрами, образующими вышеупомянутые группы.

Для определения искомого  $T_p$  воспользуемся графическими моделями, отражающими существенные аспекты периодичности. Простейшая из них показана на рис. 1.

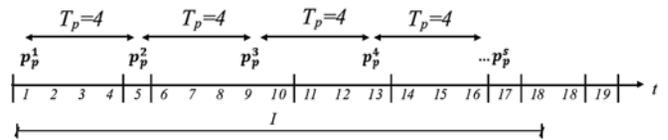


Рис. 1. Сеть единственного циклического процесса

Клоны  $p_p$  образуют конечную последовательность  $\{p_p^s\}$ , где  $s$  — порядковый номер элемента последовательности. Некоторый обозримый период времени  $I$ , в котором ведется поиск повторяющихся процессов, разбит на упорядоченные интервалы  $s$  равной длины, именуемые квантами. Каждому кванту поставлен в соответствие порядковый номер  $m$ , позволяющий как идентифицировать кванты, так и различать предшествующий и последующий кванты при их сравнении. Другими словами, имеет место последовательность  $\{c_m\}$ . Начало любого клона  $p_p$  всегда ассоциируется с конкретным  $c_m$ . Таким образом, можно говорить об отображении  $p_p \rightarrow c$ . Периодическая циклическость (далее просто циклическость) в интервале  $I$  для  $p_p$  будет иметь место в том случае, если для  $\forall \text{пар} \langle p_p^{s+1}, p_p^s \rangle$ , где  $s = \overline{1, S-1}$ , выполняется условие:  $(c_m^{s+1} - c_m^s) = T_p$ . (Здесь в соответствии с ранее введенным отображением  $c_m^{s+1} = f(p_p^{s+1})$ , а  $c_m^s = f(p_p^s)$ .)

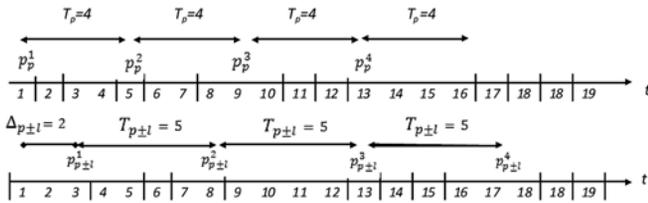


Рис. 2. Сети циклических процессов со смещением и без смещения

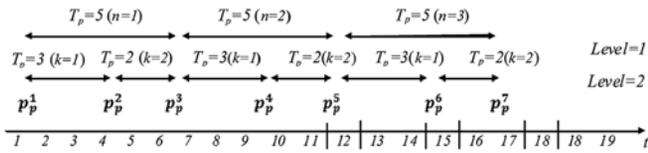


Рис. 3. Множественная (групповая) цикличность для группы из двух клонов

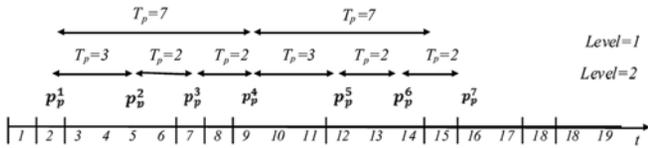


Рис. 4. Множественная цикличность для группы из трех клонов

Понятно, что начало  $p_p^1$  не всегда совпадает с  $c_1$ .

Более того, и другие  $p_p^l$ , где  $l$  — натуральное число, могут стартовать в собственные  $c_m^l$  (рис. 2). Поэтому имеет смысл для каждого  $p_p$  фиксировать смещение начала  $\Delta_p$  первого клона  $p_p^1$  от начала координат.

Рассмотренная цикличность не является единственной. Общий случай, когда повторяются группы процессов, показан на рис. 3 и 4. Период  $T_p$  для

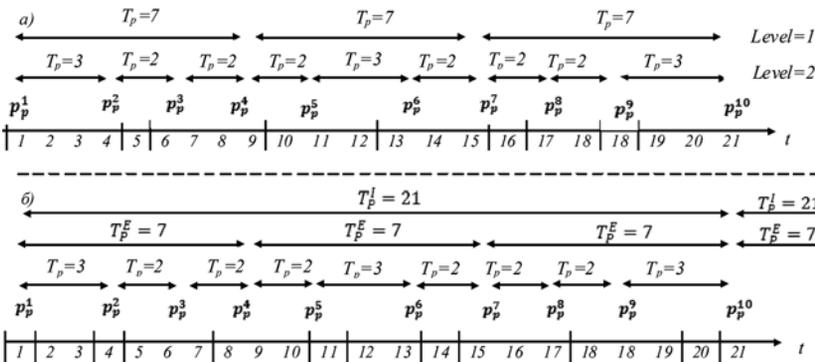


Рис. 5. Групповая цикличность

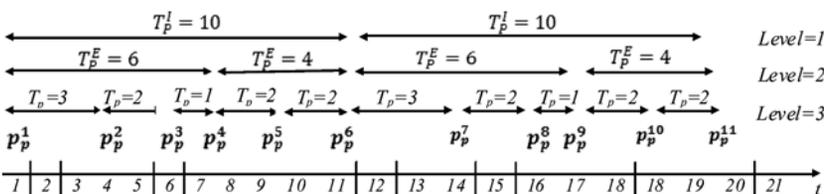


Рис. 6. "Шаговая" периодическая цикличность

группы из двух процессов (рис. 3) составляет 5 квантов, трех процессов (рис. 4) — 7 квантов.

Заметим, что для существования периодичности должно выполняться условие, несколько отличное от сформулированного ранее. Если более точно, то речь идет о комплексе условий. Применительно к ситуации, представленной на рис. 3, эти условия следующие:

$$\begin{cases} (c_5 - c_2) = (c_{10} - c_7) = (c_{15} - c_{12}); \\ (c_7 - c_5) = (c_{12} - c_{10}) = (c_{17} - c_{15}). \end{cases}$$

Для случая, показанного на рис. 4, должны выполняться условия

$$\begin{cases} (c_5 - c_2) = (c_{12} - c_9); \\ (c_7 - c_5) = (c_{14} - c_{12}); \\ (c_9 - c_7) = (c_{16} - c_{14}). \end{cases}$$

Опираясь на приведенные примеры, запишем в общем виде условие групповой цикличности, которое для  $\forall$  периодов  $n$ , содержащих одинаковое  $k$ -е число процессов, и фиксированных  $k$  имеет вид

$$f(p_p^{n(k+1)}) - f(p_p^{nk}) = \text{const}. \quad (2)$$

Нельзя оставить без внимания ситуацию, когда условие (2) не соблюдается, но группы процессов, тем не менее, повторяются с известной регулярностью (рис. 5). С периодичностью  $T_p = 7$  повторяются группы, состоящие из трех клонов.

Тем не менее, можно предположить, что при значительном увеличении продолжительности анализируемого периода  $I$  (как минимум в два раза) искомая периодичность будет обнаружена. Пример подобного рода приводится на рис. 5, б. Пусть повторяются группы из девяти клонов с интегральным (действительным) периодом  $T_p^I$ , равным 21-му кванту. При этом условие (2) выполняется и, следовательно, присутствует периодическая цикличность.

Чтобы различать периоды  $T_p^I$  и  $T_p^E$ , последний будем именовать локальным периодом повторения, а соответствующую последовательность принадлежащих ему клонов — вложенной последовательностью (по аналогии с вложенными циклами, используемыми в программировании).

В завершение рассмотрения вариантов групповой цикличности приведем пример вложенной последовательности, элементы которой повторяются с некоторым шагом (смещением). Величина последнего в общем случае определяется как  $\sum_{j=1 \div N} T_p^j$ , где  $j$  — номер последовательности, следующей за искомой (рис. 6). Здесь последовательность из трех клонов повторяется с шагом  $\sum_{j=1 \div 1} T_p^j = 4$ .

Можно предположить, что существует иерархия групп со сколь угодно большим числом уровней иерархий *Level*, каждый из которых может включать вложенные последовательности как с нулевым шагом, так и с шагом, отличным от нуля. Но это предположение легко опровергается, поскольку все промежуточные уровни — это не более чем манипуляции с разбиением одного единственного базового интегрального уровня, содержащего периоды  $T_p^I$ , на подуровни.

Кроме того, подобная декомпозиция не оказывает никакого влияния на решение поставленной ранее задачи по полному и компактному представлению периодических данных в базах данных, поскольку фиксации подлежат все начальные кванты клонов, образующих интегральную последовательность с периодом  $T_p^I$ .

Рассмотренные комбинации периодической цикличности позволяют сформировать полный набор параметров, которые должны быть представлены в базе данных. Это, кроме очевидных  $T_p$  и  $\Delta_p$ , кванты начала каждого из клонов в процессной группе.

Полученные формальные выражения (1), (2) дают возможность построить однопроходной алгоритм нахождения всех перечисленных параметров для известных последовательностей экземпляров процессов и соответствующих им квантов. (Представленная ниже программа не содержит блока формирования входного массива *Input*, так как это несущественно для рассматриваемого ниже алгоритма.)

Алгоритм последовательно считывает все  $(c_m^{s+1} - c_m^s)$  из массива *Input* и размещает их в двумерном массиве *Output* размерностью  $r \times c$ , где  $r = c = M$ . Каждое очередное значение сначала хранится в переменной *current\_value*, а затем размещается в элементе *Output*, индексируемом текущими значениями  $r$  и  $c$ . При равенстве  $Output[r, c]$  и  $Output[r - 1, c]$  изменяются индексы  $r$  или  $c$ , в противном случае, за счет пересылки на очередные свободные ячейки первой строки данных либо из первого столбца (если  $r > 1$ ) или (при  $r = 1$ ) данных второй строки — формируется последовательность процессов, образующих циклическую группу.

На основании значений, принимаемых переменными  $i$  и  $r$  после завершения работы алгоритма, можно сделать вывод о наличии или отсутствии периодичности, а также получить искомые параметры цикла, если таковой присутствует в анализируемой последовательности, состоящей из  $(c_m^{s+1} - c_m^s)$ . Если  $r > 1$ , то наблюдается цикличность с числом циклов, равным  $r$ . Значение, которое содержится в  $i$ , — не что иное как число элементов (клонов) в циклической группе.

Листинг 1 представляет текст программы на языке C#, реализующий представленный алгоритм.

**Листинг 1.** Расчет параметров цикличности  
 int r = 1, c = 0, i = 1; // r и c — текущие индексы соответственно строки и столбца

```
bool first = true; // массива Output
int[] Input = new int[k]; // i — индекс текущего элемента массива Input
int[,] Output = new int[k, k];
foreach (int current_value in Input)
{ if (first) { Output[0, 0] = Input[0];
  first = false;
} else { Output[r, c] = current_value;
  if (Output[r, c] == Output[r - 1, c])
  { if ((i - c) > 1) c++;
    else { r++; c = 0; } } else { if (r > 1) // Пересылка столбца
    { for (int l = 1; l < r + 1; l++, i++)
      Output[0, l] = Output[l, 0]; r = 1;
    } else // Пересылка строки
    { for (; c >= 0; c--, i++) Output[0, i] = Output[l, c];
      c = 0; } } } }
```

При последующем структурном моделировании будем обращать отдельное внимание на случай, когда для какой-то части процессов регулярность (периодичность) отсутствует. Очевидно, что сочетание периодических и непериодических последовательностей процессов также должно быть особым образом отражено на уровне структур данных.

### Моделирующие структуры

В настоящем разделе рассматриваются, анализируются и сравниваются "циклические" и "нециклические" моделирующие структуры, которые могут быть привлечены для размещения в них повторяющихся пар  $\langle e, p \rangle$  соответственно периодических  $pd$  и непериодических  $nd$  данных. В "нециклических" отношениях, в отличие от "циклических", хранятся все клоны экземпляров  $p_p$ . Факт повторяемости клонов в расчет не принимается ввиду отсутствия в их последовательностях постоянного периода повторения.

Типичная "нециклическая" структура в составе кластера, куда вошли также все "родительские" объекты, показана на рис. 7. Непосредственно сами данные о повторяющихся процессах размещаются в *Trivial*-структуре. Порядковые номера квантов, как и, возможно, их продолжительности, если длины всех квантов не равны между собой, хранятся

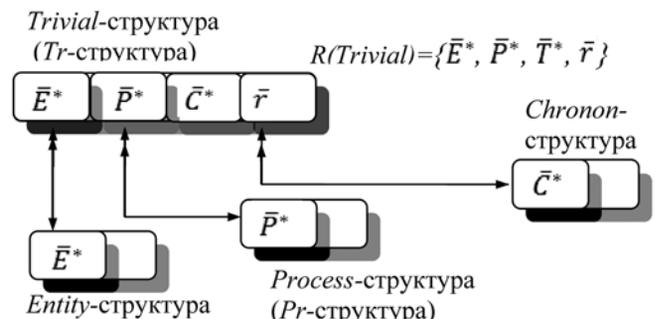


Рис. 7. Организация "нециклического" кластера

в *Chronon*-структуре. Экземпляры сущностей и процессов содержатся соответственно в *Entity*- и *Process*-структурах. (Далее, на части последующих диаграмм, если не требуется раскрывать атрибутивный состав, *Entity*- и *Process*-структуры, включая и связи, ведущие к ним, будут опускаться.)

Универсальная, многоцелевая "циклическая" конструкция представлена на рис. 8. Ее атрибутивный состав позволяет отразить все установленные ранее параметры повторяемости. Обратим особое внимание на атрибут  $\bar{T}_p$  *Entity*-структуры, тем более, что это единственное, что отличает "нециклический" кластер (см. рис. 7) от кластера, в котором учтена повторяемость процессов (см. рис. 8). Соответствующие  $\bar{T}_p$  элементы данных содержат информацию о  $T_p$ . Если более точно — то о длине периода, которая измеряется в квантах, для клонов экземпляра  $p_p$ .

$T_p$  — это расчетная величина и она может быть получена из *Multipurpose*-отношения, если выполняется одно условие. Оно сводится к тому, что в *Multipurpose*-отношении для каждого  $p_p$  должен присутствовать дополнительный кортеж, позволяющий вычислить  $T_p$ . Два примера реляционных отношений, иллюстрирующих сказанное, с использованием атрибута  $T_p$  и без него, для "циклического" кластера продемонстрированы на рис. 9.

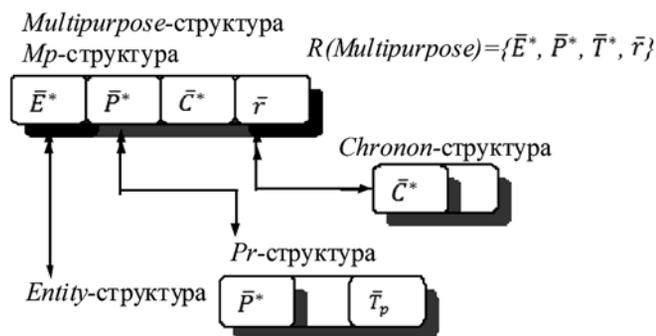


Рис. 8. Структурная организация "циклического" кластера

а) <i>Mp</i> -структура (группа из трех процессов)						<i>Mp</i> -структура (группа из одного процесса)					
$\bar{E}^*$	$\bar{P}^*$	$\bar{C}^*$	$\bar{r}$			$\bar{E}^*$	$\bar{P}^*$	$\bar{C}^*$	$\bar{r}$		
1	1	1	1	$t_1$		1	1	2	1	$t_1$	
1	1	4	1	$t_2$		1	1	4	1	$t_2$	
1	1	6	1	$t_2$							
1	1	8	1	$t_4$							
$T_{p=1} = t_4(\bar{C}) - t_1(\bar{C}) = 7$						$T_{p=1} = t_2(\bar{C}) - t_1(\bar{C}) = 2$					
б) <i>Mp</i> -структура						<i>Mp</i> -структура					
$\bar{E}^*$	$\bar{P}^*$	$\bar{C}^*$	$\bar{r}$			$\bar{E}^*$	$\bar{P}^*$	$\bar{C}^*$	$\bar{r}$		
1	1	1	1	$t_1$		1	1	2	1	$t_1$	
1	1	4	1	$t_2$							
1	1	6	1	$t_2$							
<i>Ps</i> -структура						<i>Ps</i> -структура					
$\bar{P}^*$	$\bar{T}_p$					$\bar{P}^*$	$\bar{T}_p$				
1	7					1	2				

Рис. 9. Примеры альтернативных реляционных отношений, учитывающих циклическость данных

Кортеж  $t_4$  для группы из трех процессов и кортеж  $t_2$  для группы, включающей один процесс, выполняют функцию дополнительного кортежа. На основании данных, которые в них содержатся, можно рассчитать  $T_p$ . Для этого достаточно для каждого экземпляра  $p_p$  провести поиск клона, стартовавшего раньше других, и далее воспользоваться соотношением

$$T_{p=d} = t_{p=d}^m(\bar{C}) - t_{p=d}^n(\bar{C}), \quad (3)$$

где  $m = \max\{t_{p=d}(\bar{C})\}$ ,  $n = \min\{t_{p=d}(\bar{C})\}$ .

Альтернативная конструкция, не использующая дополнительный кортеж, показана на том же рис. 9, б. Значения  $T_p$  заранее рассчитаны и размещены в *Ps*-структуре.

Применение "циклических" отношений сопровождается эффектом компактности. Получим оценку отношения мощностей  $M_p$ - и  $T_r$ -структур:

$$\frac{M(M_p)}{M(T_r)} = \frac{\sum_{p=1}^P p_p}{P \sum_{p=1}^P s_p}, \quad (4)$$

где  $p_p$  — группа процессов;  $s_p$  — порядковый номер группы повторяющихся процессов;  $P$  — число повторяющихся процессов;  $S_p$  — число повторений  $p_p$ .

Очевидно, что с ростом числа повторений растет и отношение (4), следовательно, увеличивается компактность  $M_p$ -структуры.

Исследуем выражение (4), представив его в виде функции  $\frac{M(M_p)}{M(T_r)} = M = f(p, s)$ . Зафиксируем пред-

варительно  $p$ . Тогда  $f(p, s)$  при  $p = \text{const}$  — это гиперболоа, которая достаточно быстро убывает с ростом  $s$ , что свидетельствует о предпочтении использования *Mp*-структуры в сравнении с *Tr*-структурой.

Отметим, что ссылка на квант, идентифицирующий начало процесса (атрибут  $\bar{C}^*$ ), делает ненужным специальным образом фиксировать смещение  $\Delta_p$ . Правда, при этом следует придерживаться соглашения, согласно которому область определения (доменом) атрибута *Chronon* ( $\bar{C}^*$ ) является множество натуральных чисел, что соответствует характеристике введенной ранее последовательности  $\{c_m\}$ .

### Встраивание неперiodических циклических данных в циклические структуры

Ситуация, когда отдельные экземпляры пары "тип сущности — тип процесса" не содержат  $pd$ , тогда как большинство других такие данные включают, не является исключительной. Задача совместного размещения *периодических* и *непериодических* повторяющихся данных имеет простое решение.

Идентичность атрибутивного состава *Tr*- и *Mp*-структур позволяет сосредоточить одновременно периодические и непериодические данные в одной

Mp-структура					Ps-структура		
$\bar{E}^*$	$\bar{P}^*$	$\bar{C}^*$	$\bar{r}$		$\bar{P}^*$	$\bar{T}_p$	
1	1	2	1	$t_1$	1	7	$t_1$
1	1	4	1	$t_2$	2	3	$t_2$
1	1	6	1	$t_2$	3	.null	$t_3$
1	2	9	1	$t_4$	Ch-структура		
1	3	7	1	$t_5$	$\bar{C}^*$		
1	3	9	1	$t_6$	1		$t_1$
1	3	15	1	$t_7$	2		$t_2$
					...		
					15		$t_{15}$
					...		

Рис. 10. Использование "нулевых" кортежей для представления неперiodических процессов

единственной структуре. Речь, в частности, идет об *Mp*-структуре. При этом, правда, требуется каким-то образом различать периодические и неперiodические наборы данных. Вариант с использованием дополнительного кортежа *Mp*-отношения не подходит (рис. 9), поскольку неперiodические данные в этом случае могут интерпретироваться как периодические, и наоборот, периодические восприниматься как неперiodические. В то же время использование нулевых (.null) или "пустых" значений, которые может принимать атрибут  $\bar{T}_p$  *Ps*-структуры, является вполне приемлемым решением. Достаточно условиться, что если атрибут  $\bar{T}_p$  какого-либо кортежа *t* *Ps*-структуры равен .null ( $t(\bar{T}_p) = .null$ ), то этот кортеж указывает на неперiodический процесс. Пример, приведенный на рис. 10, наглядно демонстрирует сказанное. Таким кортежем, описывающим и идентифицирующим неперiodический процесс, является кортеж  $t_3$  *Ps*-структуры.

Тем самым, одновременно достигаются две цели: компактное размещение периодических данных и единообразное (в одних и тех же структурах) хранение периодических и неперiodических наборов.

### Привязка прогнозируемых траекторий к реальному времени

Задача привязки к оси реального времени на проверку оказывается не такой тривиальной, как это может показаться. На первый взгляд, достаточно из всех "первых" процессов в парах  $\langle e, p \rangle$  выбрать процесс, стартующий раньше других, и задать время его начала. При более тщательном рассмотрении оказывается, что не исключена вероятность того, что, например, со временем может появиться процесс, начинающийся раньше выбранного. И это только одна из потенциальных нештатных ситуаций. Их все можно исключить, если поставить в соответствие метку времени начала не какому-то отдельному экземпляру процессов, а всему типу  $\langle e, p \rangle$ .

Принимая во внимание существование множества пар  $\langle e, p \rangle$ , в ту часть модели данных, которая обычно отводится под размещение разнообразной метаинформации, логично включить структуру со следующей схемой:

$$R(\text{TimeTuning}) = \{Id\_ob^*, \text{TimeStamp}, Id\_mnc\},$$

где *Id\_ob* — внутренний идентификатор (используемый в конкретной модели данных) *Mp*-структуры; *TimeStamp* — метка времени, синхронизированная с квантами, размещенными в родительской по отношению к *Mp* *Ch*-структуре.

Под синхронизацией здесь подразумевается использование одних и тех же единиц измерения времени, что для квантов, что для меток времени. Если, например, кванты измеряются в днях, то и *TimeStamp* должен принимать значение с точностью до дня. Атрибут *Id\_mnc* в отношении со схемой *R(TimeTuning)* есть не что иное как ссылка на используемую шкалу времени.

Несмотря на то что предложенный вариант привязки носит системный характер, поскольку охватывает все типы пар  $\langle e, p \rangle$ , он пока никак не связан с тем, что можно было бы назвать моделью реального времени базы данных. Построение такой модели — предмет отдельного самостоятельного исследования. В связи с этим уместно сделать ссылку на ряд работ в области темпоральных баз данных [3, 5, 11], авторы которых разрабатывают составные части подобной временной модели. Полученные ими результаты пока охватывают только применяемые временные категории и различные расширения *Entity-Relationship*-технологии в части отражения различных временных аспектов представления сущностей и их взаимодействий в реальной среде.

Указанное обстоятельство делает неполным предложенный вариант привязки, хотя он и включает в себя все наиболее значимые аспекты, в том числе и закрывает вопрос с отражением в моделях данных последнего, не затронутого ранее параметра цикличности — смещения начала первого процесса  $\Delta_p$ .

### Заключение

В работе нашли отражение результаты исследований, касающихся поиска источников, идентификации, моделирования и представления в базах данных периодических данных организационных систем. Показано, что такие данные, как правило, используются в прогнозируемых траекториях и всегда ассоциируются с парой  $\langle \text{сущность}, \text{процесс} \rangle$ . Упомянутая локализация — следствие естественного стремления организационных систем к ритмичному функционированию, которое явно просматривается в фазах планирования.

Присутствие цикличности делает возможным не фиксировать в базах данных всю траекторию, а ограничиться только указанием характеристики цикла. С этой целью в работе построены модели

периодичности и на их основе разработан алгоритм нахождения характеристик периодичности для циклов, состоящих как из одного, так и из нескольких процессов. Предложены моделирующие, в том числе и альтернативные, структуры, позволяющие организовать хранение циклических характеристик в базе данных.

За рамками работы остались несколько вопросов, которые могут продолжить начатое исследование — это встраивание "циклических структур" в подсистему единого модельного времени базы данных (которую еще нужно создать), и подбор моделирующих структур для декомпозируемых и агрегируемых периодических данных с формированием системы соответствующих ограничений. Последняя задача имеет ряд серьезных приложений для функциональной составляющей информационных систем. В частности, это касается нахождения подмножеств зависимых данных, которые должны быть подвержены изменениям вследствие корректировки данных на одном из уровней агрегирования. Функциональные задачи могут охватывать как задачи модификации еще не выполняемых, так и уже частично реализуемых прогнозируемых траекторий.

1. Spaccapietra S., Parent C., Damiani M., Macedo J., Porto F., Vangenot C. A conceptual view on trajectories // *Data and knowledge engineering*. 2008. Vol. 65. P. 126—146.
2. Egidì L., Terenziani P. A modular approach to user-defined symbolic periodicities // *Data and knowledge engineering*. 2008. Vol. 66. P. 163—198.
3. Date C. J., Hugh Darwen, Lorentzos N. A. *Time and relation theory*. Morgan Kaufmann Publishers, 2014. 560 p.
4. Jimenez L. J. REEEM: Reenhancing the entity-relationship model // *Data and knowledge engineering*. 2006. Vol. 58. P. 410—435.
5. Gregerson H., Jensen C. S. Temporal entity-relationship models — a survey // *IEEE Transaction on Knowledge and Data Engineering*. 1999. Vol. 11 (3). P. 464—497.
6. Джексон Г. Проектирование реляционных баз данных для использования с микроЭВМ. М.: Мир, 1991. 252 с.
7. Мейер Д. Теория реляционных баз данных. М.: Мир, 1987. 608 с.
8. Moody D. L. Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions // *Data and knowledge engineering*. 2005. Vol. 55 (3). P. 243—276.
9. Родионов А. Н. Качество даталогических схем. Принцип компактности моделей данных и его приложения I // *Информатика и системы управления*. 2012. № 1. С. 16—27.
10. Мартин Дж. Организация баз данных в вычислительных системах. М.: Мир, 1980. 663 с.
11. Jensen C. S., Dyreson C. E. (Eds). The consensus glossary of temporal database concepts // *Temporal Databases: Research and Practice*. Springer Verlag. 1998. P. 367—405.

A. N. Rodionov, Leading Researcher, e-mail: ran@newmail.ru,  
Computer Centre of Far-Eastern Branch of RAS

## Periodical Data: Identification, Modeling and Representation in Databases

*There are some tasks concerned with the modeling of such data that differs by the specific properties. This is a data repeated periodically (a cycle data) and a data repeated no periodically. Its representation in databases must capture the separate actions being a parts of entities' lifespan. To that end it is necessary to implement system decision that comprises the temporal and process aspects.*

*In this article we examine the full range of the relevant issues. Thirsts of all the identification characteristics and the cycle attributes are designated. Then some models for finding parameters of cycle processes are constituted. Based on the derived model it introduces and researches the modeling constructions after which compare their quantity. The relations corresponding to modeling structures may utilize to accommodate both cycle and joint data. We also proposal and substantiate the multipurpose structure to store periodical and not periodical data simultaneously.*

*One of the unexpected results appeared to be the fact that the cyclic and acyclic structures, each of which is designed for storage recurrent periodical and recurrent not periodical data, are fully identical. Finally, it articulates how to connect the modeling structures to time axis of database and specifies the advanced fields of study.*

**Keywords:** *periodical and not periodical cycle data, type, instance and clone of entity, multipurpose data structures, temporal databases*

### References

1. Spaccapietra S., Parent C., Damiani M., Macedo J. Porto F., Vangenot C. A conceptual view on trajectories. *Data and knowledge engineering*. 2008, vol. 65, pp. 126—146.
2. Egidì L., Terenziani P. A modular approach to user-defined symbolic periodicities. *Data and knowledge engineering*. 2008, vol. 66, pp. 163—198.
3. Date C. J., Hugh Darwen, Nikos A. Lorentzos. *Time and relation theory*. Morgan Kaufmann Publishers, 2014. 560 p.
4. Jimenez L. J. REEEM: Reenhancing the entity-relationship model. *Data and knowledge engineering*. 2006, vol. 58, pp. 416—435.
5. Gregerson H., Jensen C. S. Temporal entity-relationship models — a survey. *IEEE Transaction on Knowledge and Data Engineering*. 1999, vol. 11 (3), pp. 464—497.
6. Dzheksion G. *Proektirovanie reljacionnyh baz dannyh dlja ispol'zovanija s mikro-JeVM*. М.: Мир, 1991. 252 p.
7. Mejer D. *Teorija reljacionnyh baz dannyh*. М.: Мир, 1987. 608 p.
8. Moody D. L. Theoretical and practical issues in evaluating the quality of conceptual models: current state and future directions. *Data and knowledge engineering*. 2005, vol. 55 (3), pp. 243—276.
9. Rodionov A. N. Kachestvo datalogicheskikh shem. Princip kompaktnosti modelej dannyh i ego prilozhenija I. *Informatika i sistemy upravlenija*. 2012, no. 1, pp. 16—27.
10. Martin Dzh. *Organizacija baz dannyh v vychislitel'nyh sistemah*. М.: Мир, 1980. 663 p.
11. Jensen C. S., Dyreson C. E. (Eds). The consensus glossary of temporal database concepts. *Temporal Databases: Research and Practice*. Springer Verlag. 1998, pp. 367—405.

**Я. В. Катуева**, канд. техн. наук, ст. науч. сотр., gloria@iacp.dvo.ru,  
**Д. А. Назаров**, канд. техн. наук, науч. сотр., nazardim@iacp.dvo.ru  
 Институт автоматизации и процессов управления ДВО РАН, г. Владивосток

## Методы параметрического синтеза на основе сеточного представления области работоспособности

*Обсуждается применение методов параметрического синтеза для обеспечения параметрической надежности с использованием сеточного приближения области работоспособности исследуемой системы. Предложена реализация аналитического метода механической и электрической аналогий для дискретного представления области. Для метода максимального запаса работоспособности предложены различные способы вычисления целевой функции, основанные на вписании в область выпуклых симметричных фигур максимального объема. Реализован метод сужающихся областей для дискретного сужения границы.*

**Ключевые слова:** качество, надежность, работоспособность, параметрический синтез, область работоспособности

### Введение

Проектирование аналоговых технических систем на этапе параметрического синтеза сводится к определению номинальных значений внутренних (первичных, управляемых) параметров синтезируемого объекта (исследуемой системы), при которых будут удовлетворены условия технического задания. При оптимальном параметрическом синтезе технических устройств для обеспечения параметрической надежности необходимым условием является учет процессов деградации параметров под воздействием различных дестабилизирующих факторов, в том числе и производственно-технологического характера.

К сожалению, априорные данные о вероятностных характеристиках деградаций параметров или факторов (возмущений), приводящих к ним, и, как следствие, деградационным (параметрическим) отказам, зачастую неизвестны, поэтому приходится принимать рационально обоснованные решения в ходе проектирования систем при отсутствии полной информации. Для этого необходимо сделать ее устойчивой к максимальному подмножеству возмущений, ориентируясь на наихудший случай. Тогда в качестве целевой функции оптимального параметрического синтеза для обеспечения параметрической надежности предлагается использовать запас работоспособности [1–3].

Для нахождения значений номиналов параметров, имеющих максимальный запас работоспособности, предлагаются геометрические методы, основанные на предположении, что известно сеточное представление области работоспособности — ее дискретная аппроксимация, построенная с помощью регулярной сетки и представленная в виде соответствующей структуры [4].

### Задача параметрического синтеза по критерию запаса работоспособности

Пусть определена структура исследуемого объекта (системы) и в результате дополнительных исследований определены внутренние варьируемые доминирующие параметры  $\mathbf{x} = (x_1, \dots, x_n)^T$ , которые характеризуют состояние и свойства системы, при этом статистические данные о законах распределения процессов их деградации, в том числе обусловленной производственно-технологическими факторами, неизвестны. Параметры  $\mathbf{x} = (x_1, \dots, x_n)^T$  в работе будем называть внутренними параметрами.

Для исследуемой системы построена математическая модель — функциональная зависимость выходных параметров  $\mathbf{y} = (y_1, \dots, y_m)^T$  от параметров элементов  $\mathbf{x}$ :

$$y_j = y_j(x_1, \dots, x_n). \quad (1)$$

Возможные вариации значений внутренних параметров задаются из условий их физической реализуемости, ограничений технического задания и представляют собой брус допусков

$$B_T = \{x_{i\min} \leq x_i \leq x_{i\max}, i = \overline{1, n}\}. \quad (2)$$

Показателем функционирования системы является выполнение условий работоспособности:

$$\mathbf{y}_{\min} \leq \mathbf{y}(\mathbf{x}) \leq \mathbf{y}_{\max}. \quad (3)$$

Условия (3) определяют в пространстве внутренних параметров область работоспособности  $D_x$ :

$$D_x = \{\mathbf{x} \in R^n | \mathbf{y}_{\min} \leq \mathbf{y}(\mathbf{x}) \leq \mathbf{y}_{\max}\}. \quad (4)$$

Поскольку основным свойством объекта исследования является отсутствие статистически достоверной информации о законах деградации параметров, предлагается использовать в качестве целевой функции процедуры оптимального параметрического синтеза запас работоспособности на уровне внутренних параметров системы [2]. Он позволяет

оценить степень удаленности вектора внутренних параметров  $\mathbf{x} = (x_1, \dots, x_n)^T$  от границы области работоспособности  $\partial D_x$ , а следовательно, пределы возможных вариаций параметров элементов, при которых не нарушаются условия работоспособности. Задача оптимального параметрического синтеза в этом случае сводится к нахождению такой точки  $\mathbf{x}_{\text{opt}}$  внутри области работоспособности  $D_x$ , которая находится на максимальном расстоянии от ее границы  $\partial D_x$  [1, 2, 5]:

$$\mathbf{x}_{\text{opt}} = \arg \max_{\mathbf{x} \in D_x \cap B_T} \text{dist}(\mathbf{x}, \partial D_x), \quad (5)$$

где  $\text{dist}(\mathbf{x}, \partial D_x)$  — каким-либо способом измеренное минимальное расстояние от точки  $\mathbf{x} = (x_1, \dots, x_n)^T$  до границы области работоспособности, соответствующее запасу работоспособности в точке  $\mathbf{x} = (x_1, \dots, x_n)^T$ .

Отметим, что область работоспособности  $D_x$  и ее граница  $\partial D_x$  также неизвестны. Одним из способов аппроксимации неизвестной многомерной области является построение ее дискретного аналога — сеточного представления области работоспособности, в основе которого лежит метод матричных испытаний на надежность [4].

В работе рассматриваются методы решения задачи оптимального параметрического синтеза (5), использующие сеточное (дискретное) представление области работоспособности, когда эта область представлена в виде множества непересекающихся  $n$ -мерных параллелепипедов с использованием различных способов формирования целевой функции.

### Аппроксимация области работоспособности на основе регулярной сетки

Будем считать, что с помощью регулярной сетки  $B_0^g$  в описанном брусе  $B_0$  с границами  $a^0, b^0$  и числом узлов по каждому координатному направлению  $l_i, i = \overline{1, n}$ , построен дискретный аналог области работоспособности  $D_x^0$  [4]. Общее число узлов сетки  $L = \prod_{i=1}^n l_i$ .

Шаг сетки по каждому из внутренних параметров составляет

$$h_i = (b_i^0 - a_i^0) / l_i, \quad i = \overline{1, n}. \quad (6)$$

Описанный брус является объединением брусов-ячеек:

$$B_0 = B_0^g = \bigcup_{k_1=1}^{l_1} \bigcup_{k_2=1}^{l_2} \dots \bigcup_{k_n=1}^{l_n} e_{k_1, k_2, \dots, k_n}, \quad (7)$$

где

$$e_{k_1, k_2, \dots, k_n} = \{\mathbf{x} \in B_0 | a_i^0 + (k_i - 1)h_i \leq x_i \leq a_i^0 + k_i h_i, \forall i = \overline{1, n}\}.$$

Каждый из брусов-ячеек  $e_{k_1, k_2, \dots, k_n}$  (узлов сетки) однозначно определяется соответствующим набором индексов  $(k_1, k_2, \dots, k_n), 1 \leq k_i \leq l_i, \forall i = \overline{1, n}$ .

Для узла сетки  $e_{k_1, k_2, \dots, k_n}$  координаты точки-представителя  $\mathbf{x}^g$ :

$$x_i^g = a_i^0 + k_i h_i - h_i / 2, \quad \forall i = \overline{1, n}. \quad (8)$$

Введем вспомогательную функцию

$$\chi(\mathbf{x}^g) = \begin{cases} 1, & \text{если } \mathbf{y}_{\min} \leq \mathbf{y}(\mathbf{x}^g) \leq \mathbf{y}_{\max}, \\ 0, & \text{в противном случае,} \end{cases} \quad (9)$$

являющуюся характеристической функцией дискретизации области работоспособности  $D_x$ .

Множество всех узлов сетки, для которых выполняются условия работоспособности (3),

$$D_x^0 = \{e_{k_1, k_2, \dots, k_n} \in B_0 | \chi(\mathbf{x}^g(k_1, k_2, \dots, k_n)) = 1 \forall k_i = \overline{1, l_i}, i = \overline{1, n}\} \quad (10)$$

является дискретным или сеточным приближением  $D_x$ . Число элементов сетки, попавших в множество  $D_x^0$ , обозначим  $M = |D_x^0|$ .

Для хранения данных о сетке  $B_0^g$  и множестве  $D_x^0$  используются соответствующие специальные оригинальные структуры данных, описанные в работах [6, 7]. В ходе построения множества  $D_x^0$  может быть вычислен центр тяжести  $\mathbf{x}^c = \{x_1^c, x_2^c, \dots, x_n^c\}^T$  фигуры  $D_x^0$ , аппроксимирующей области работоспособности с координатами

$$x_i^c = \frac{1}{M} \sum_{k=1}^M x_{ki}^g, \quad i = \overline{1, n},$$

где  $\mathbf{x}^g$  — координаты соответствующей точки-представителя (узла сетки), попавшей в множество  $D_x^0$  [8]. Для выпуклых областей работоспособности центр тяжести области является решением задачи оптимального параметрического синтеза [3].

### Использование аналитического метода механической и электрической аналогий для сеточного представления области

Для решения задачи (5) воспользуемся предложенным в работах [5, 9] аналитическим методом механической и электрической аналогий для выпуклых областей работоспособности, который заключается в следующем.

Пусть задано множество  $\partial D_x^0$  точек области работоспособности:

$$\partial D_x^0 = \{X_k \in \partial D_x, k = 1, N\}, \quad (11)$$

принадлежащих границе, которые будем называть граничными точками области. Тогда решением задачи оптимального параметрического синтеза по

критерию максимального запаса работоспособности будет точка  $\mathbf{x}^0 = (x_1^0, x_2^0, \dots, x_n^0)^T$ , с координатами:

$$x_i^0 = \frac{1}{N} \sum_{k=1}^N X_{ki}, \quad i = \overline{1, n}. \quad (12)$$

Отметим, что точка  $\mathbf{x}^0 = (x_1^0, x_2^0, \dots, x_n^0)^T$  является центроидом множества точек  $\partial D_x^0$ , принадлежащих границе области работоспособности.

Для построения множества  $\partial D_x^0$  граничных элементов сетки введем понятие сопряженных (соседних) элементов. Элементы  $e_{k_1, k_2, \dots, k_n}$  и  $e_{m_1, m_2, \dots, m_n}$  сетки являются соседними (сопряженными), если для их индексов выполняется условие

$$\sum_{i=1}^n |m_i - k_i| = 1. \quad (13)$$

Пример соседних элементов сетки для двумерного случая представлен на рис. 1.

Определим множество граней сетки  $B_0^g$  аналогично определению граней куба в работе [7]. Правой гранью  $GR_i^+(B_0^g)$  сетки  $B_0^g$  в  $i$ -м координатном направлении будем называть подмножество всех элементов сетки  $B_0^g$ , для которых  $k_i = l_i$ :

$$\begin{aligned} GR_i^+(B_0^g) &= \\ &= \bigcup_{k_1=1}^{l_1} \dots \bigcup_{k_{i-1}=1}^{l_{i-1}} \bigcup_{k_{i+1}=1}^{l_{i+1}} \dots \bigcup_{k_n=1}^{l_n} e_{k_1 \dots k_{i-1}, l_i, k_{i+1} \dots k_n}. \end{aligned}$$

Левой гранью  $GR_i^-(B_0^g)$  сетки  $B_0^g$  в  $i$ -м координатном направлении будем называть, соответственно, подмножество элементов сетки, для которых  $k_i = 1$ :

$$\begin{aligned} GR_i^-(B_0^g) &= \\ &= \bigcup_{k_1=1}^{l_1} \dots \bigcup_{k_{i-1}=1}^{l_{i-1}} \bigcup_{k_{i+1}=1}^{l_{i+1}} \dots \bigcup_{k_n=1}^{l_n} e_{k_1 \dots k_{i-1}, 1, k_{i+1} \dots k_n}. \end{aligned}$$

Границей  $GR(B_0^g)$  сетки  $B_0^g$  будет множество левых и правых граней  $B_0^g$  по всем координатным направлениям:

$$GR(B_0^g) = \bigcup_{i=1}^n (GR_i^-(B_0^g) \cup GR_i^+(B_0^g)). \quad (14)$$

Элемент  $e_{k_1, k_2, \dots, k_n}$  множества  $D_x^0$  является граничным, если он принадлежит границе сетки  $GR(B_0^g)$  или хотя бы один из сопряженных с ним элементов не принадлежит области работоспособности.

Если область работоспособности построена с помощью регулярной сетки, к множеству граничных точек области работоспособности  $\partial D_x^0$  можно отнести следующие узлы:

$$\begin{aligned} \partial D_x^0 &= \{e_{k_1 k_2 \dots k_n} \in D_x^0 | e_{k_1 k_2 \dots k_n} \in \\ &\in GR(B_0^g) \vee \exists e_{m_1 m_2 \dots m_n} (k_1 \dots k_n) \notin D_x^0\}, \end{aligned} \quad (15)$$

где  $e_{m_1 m_2 \dots m_n} (k_1 \dots k_n)$  — какой-либо из множества сопряженных с  $e_{k_1 k_2 \dots k_n}$  элементов (рис. 2).

Тогда для решения задачи параметрического синтеза аналитическим методом механической и электрической аналогий (12), необходимо построить границу области работоспособности — множество  $\partial D_x^0 \subset D_x^0$  и вычислить координаты искомой оптимальной точки (12) аналогично формулам вычисления центра тяжести фигуры  $D_x^0$  [8]. Построить множество граничных элементов области работоспособности  $\partial D_x^0 \subset D_x^0$  можно, перебрав все элементы множества  $D_x^0$ , для каждого из них провести проверку выполнения условий принадлежности границе (15).

Использование для нахождения граничных точек области работоспособности ее сеточного представления решает задачу определения граничных точек, обеспечивает их равномерное расположение и возможность определения граничных точек многомерных областей любой формы.

Существенным недостатком метода механической и электрической аналогий является априорное допущение о выпуклости области работоспособности. Вычисление координат точки-центроида граничных точек (12), в отличие от нахождения

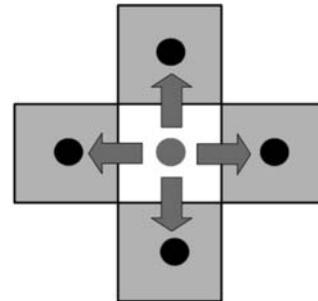


Рис. 1. Сопряженные элементы сетки в двумерном случае

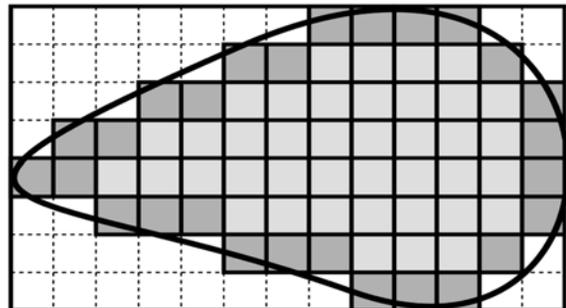


Рис. 2. Граничные элементы области в двумерном случае

центра тяжести всей области работоспособности, не может быть проведено одновременно с построением сеточного представления области  $D_x^0$ , так как для проверки принадлежности узла сетки границе области необходимо проверить все сопряженные с ним элементы. С точки зрения вычислительных затрат применение в качестве решения задачи оптимального параметрического синтеза точки — центра тяжести всей области предпочтительнее, так как вычисление ее координат может быть проведено одновременно с построением области, и кроме того, обе эти процедуры можно выполнять с использованием параллельных и распределенных вычислений на независимых параллельных процессах [10].

### Выбор оптимальных элементов сетки по критерию запаса работоспособности

Реализация критерия максимального запаса работоспособности в виде (5) подразумевает процедуру вычисления минимального расстояния  $r$  до границы для каждого из элементов  $e_{k_1, k_2, \dots, k_n} \in D_x^0$  сетки, кроме тех, которые принадлежат грани параллелепипеда  $B_0^g$  ( $k_i \neq 1, k_i \neq l_i, \forall i = \overline{1, n}$ ), присвоения этому элементу соответствующего веса  $w(e_{k_1, k_2, \dots, k_n}) = r$ , а затем выбора элемента с максимальным весом.

С точки зрения дискретной аппроксимации области работоспособности задача (5) геометрически сводится к поиску элемента сетки, являющегося центром вписанной выпуклой симметричной фигуры максимального объема, состоящей из элементов той же сетки. Под объемом фигуры будем понимать число составляющих ее элементов сетки.

Рассмотрим способы измерения расстояния  $r$  в числе узлов сетки (шагов дискретизации области работоспособности) от элемента сетки  $e_{k_1, k_2, \dots, k_n} \in D_x^0$  до границы  $\partial D_x^0$ . Отметим, что

$$r \leq \min_{i=\overline{1, n}} (l_i)/2,$$

где  $l_i, i = \overline{1, n}$ , — число узлов сетки по соответствующему координатному направлению. Для граничных элементов сетки  $r = 0$ .

Одним из способов вычисления расстояния до границы области является построение вокруг элемента сетки описанного куба максимального объема [7]. Принцип построения куба аналогичен построению окрестности Мура [11] или волновому алгоритму с 8 соседями и показан на рис. 3.

Для каждого из элементов сетки  $e_{k_1, k_2, \dots, k_n}$  значение  $r$  имеет естественные ограничения:

$$r \leq \min_{i=\overline{1, n}} (k_i - 1, l_i - k_i).$$

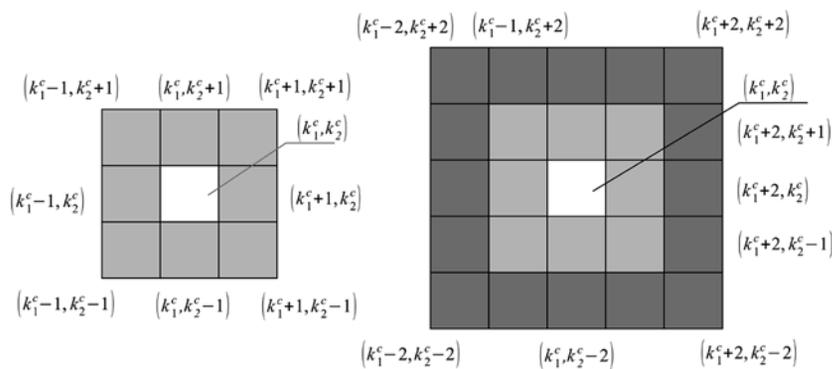


Рис. 3. Построение  $r$ -куба из элементов сетки

Множества точек, расположенных в максимальном вписанном кубе, в двумерном случае показаны на рис. 4.

Вместе с тем для вычисления минимального расстояния от элемента  $e_{k_1, k_2, \dots, k_n}$  фигуры  $D_x^0$  до ее границы  $\partial D_x^0$  можно воспользоваться понятием  $r$ -окрестности элемента сетки  $e_{k_1, k_2, \dots, k_n}$ , соответствующей окрестности фон Неймана [11] (рис. 5).

Будем называть  $r$ -окрестностью элемента сетки  $e_{k_1, k_2, \dots, k_n} \subset D_x^0$  с индексами  $(k_1, k_2, \dots, k_n)$  такое множество  $E_{k_1, k_2, \dots, k_n}^r \subset D_x^0$  элементов сетки, в котором для каждого узла  $e_{m_1, m_2, \dots, m_n} \in D_x^0$  с индексами  $(m_1, m_2, \dots, m_n)$  справедливо

$$\sum_{i=1}^n |m_i - k_i| \leq r.$$

Множество  $E_{k_1, k_2, \dots, k_n}^r$  представляет собой фигуру, состоящую из элементов сетки:

$$E_{k_1, k_2, \dots, k_n}^r = \{e_{m_1, m_2, \dots, m_n} \in B_x^g \mid \sum_{i=1}^n |m_i - k_i| \leq r, r \in \mathbf{N}\}.$$

Задача оптимального параметрического синтеза для обеспечения максимального запаса работоспособности в этом случае состоит в поиске элемента

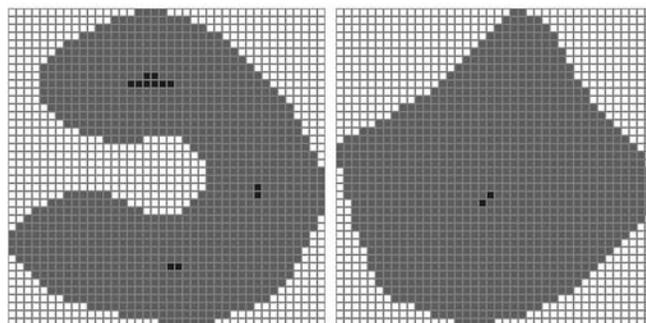


Рис. 4. Оптимальные точки области — центры максимального куба

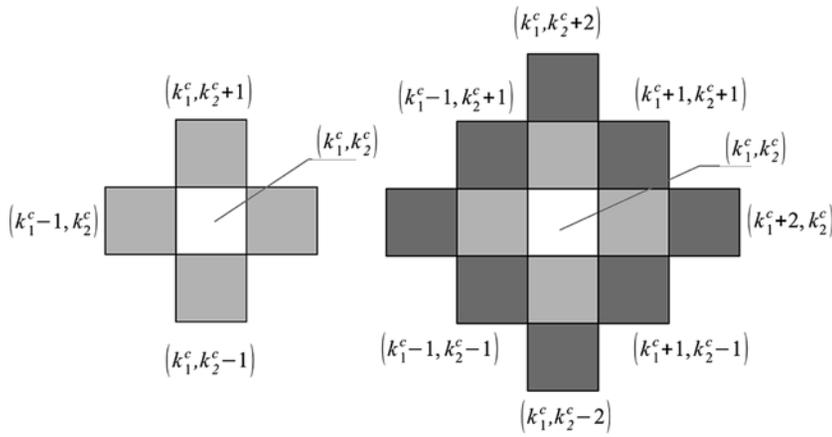


Рис. 5. Построение  $r$ -окрестности из элементов сетки

$e_{k_1^c, k_2^c, \dots, k_n^c} \in D_x^0$  сетки  $B_0^g$ , для которого можно построить максимальную  $r$ -окрестность, принадлежащую сеточному представлению области качества  $D_x^0$ .

Значение  $r$ , максимальное для данного элемента  $e_{k_1, k_2, \dots, k_n} \in D_x^0$  сеточного представления  $r$ -окрестности, показывает минимальное число шагов, за которое, двигаясь от выбранного элемента вдоль координатных направлений, можно достичь границы  $\partial D_x^0$

множества  $D_x^0$ , и соответствует манхэттенскому расстоянию от узла до границы [11]. Множество точек, расположенных на максимальном в манхэттенской метрике расстоянии от границы области и являющихся центрами максимальных вписанных окрестностей, в двумерном случае показаны на рис. 6.

Построение окрестностей фон Неймана реализуется предложенный в работах [5, 9] метод сужающихся областей для случая, когда границы области работоспособности задаются аналитически с помощью конечного числа гиперповерхностей и на каждой итерации метода происходит сужение области в сторону центра (оптимальной точки). В качестве единицы шага сужения области в этом случае

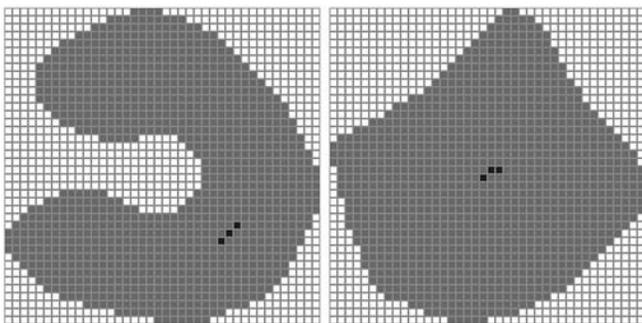


Рис. 6. Оптимальные точки области — центры максимальной окрестности

принимается элементарный брусочек (один шаг наложенной на область работоспособности сетки). Сужение границ в случае сеточного представления области работоспособности представляет собой переход к сопряженному узлу, принадлежащему области работоспособности в направлении, противоположном ближайшей границе сетки.

Граница нулевого уровня представляет собой множество  $\partial D_x^0$  граничных точек области работоспособности. Граница первого — множество точек, для которых максимальное расстояние для границы ( $r$ -окрестность) равно 1

(множество точек фигуры  $D_x^0$ , сопряженных с граничными  $\partial D_x^0$ ) и т. д. На последней итерации сужения области работоспособности будет выявлено то ее подмножество, которое составляют оптимальные в смысле критерия максимальной окрестности элементы.

Для каждого из элементов сетки может быть вычислен его "уровень сужения границы", равный минимальному числу сопряженных узлов от исследуемого узла сетки до границы области работоспособности  $\partial D_x^0$ , что соответствует максимально возможному значению  $r$  принадлежащей множеству сеточного представления  $r$ -окрестности фон Неймана элемента  $e_{k_1, k_2, \dots, k_n} \in D_x^0$ .

Использование в методе сужающихся областей сеточного представления области работоспособности, несмотря на большую вычислительную емкость метода, имеет ряд преимуществ:

- нет необходимости при каждом шаге сужения определять граничные точки;
- не нужно проводить аппроксимацию гиперповерхности, составляющей границу (полиномами, гиперсферами), и строить для них конечное число гиперповерхностей;
- не нужно строить для границ эквидистанционные поверхности и множества точек, через которые они проходят, на каждом шаге сужения области;
- нет необходимости проверять, не выродилась ли в точку область сужения;
- шаг поиска всегда задан — один шаг сетки.

В больших выпуклых областях работоспособности построение куба максимального радиуса позволит получить глобальное решение задачи оптимального параметрического синтеза; построение окрестности максимального радиуса дает множество локальных оптимумов для нахождения максимального запаса работоспособности.

Построение дискретной аппроксимации области работоспособности требует больших вычислительных затрат и ресурсов для хранения данных, при этом обеспечивает наиболее полное описание области с априори неизвестной конфигурацией, что особо актуально для сложных систем, характеризующихся большим числом варьируемых параметров. Описанные в работе методы позволяют решать задачу оптимального параметрического синтеза с учетом требований параметрической надежности в условиях отсутствия информации о тенденциях параметрического дрейфа с использованием критерия максимального запаса работоспособности. Нахождение максимальных диапазонов допустимой вариации параметров внутри этой области позволит в дальнейшем оценить чувствительность системы к изменениям отдельных параметров и выявить ее уязвимость к отдельным дестабилизирующим факторам.

Рассмотренные алгоритмы были апробированы на тестовых примерах при решении задач параметрического синтеза разнообразных аналоговых радиоэлектронных схем.

*Работа выполнена при частичной поддержке гранта РФФИ № 14-08-31291 "Эффективные методы и алгоритмы построения областей работоспособности аналоговых технических систем".*

1. **Норенков И. П., Маничев В. Б.** Основы теории и проектирования САПР: учеб. для вузов. М.: Высш. шк., 1990. 335 с.
2. **Абрамов О. В., Катueva Я. В., Назаров Д. А.** Оптимальный параметрический синтез по критерию запаса работоспособности // Проблемы управления. 2007. № 6. С. 64—69.
3. **Саушев А. В.** Параметрический синтез электротехнических устройств и систем. СПб.: ГУМРФ им. адмирала С. О. Макарова, 2013. 315 с.
4. **Катueva Я. В., Назаров Д. А.** Алгоритмы анализа области работоспособности, заданной в матричной форме // Информатика и системы управления. 2005. № 2 (10). С. 118—128.
5. **Саушев А. В.** Области работоспособности электротехнических систем. СПб.: Политехника, 2013. 414 с.
6. **Назаров Д. А.** Использование областей работоспособности для оптимального выбора номиналов параметров // Информатика и системы управления. 2011. № 2 (28). С. 59—69.
7. **Аноп М. Ф., Катueva Я. В., Назаров Д. А.** Использование области качества в задачах обеспечения параметрической надежности // Информатика и системы управления. 2014. № 2 (40). С. 66—74.
8. **Катueva Я. В., Евдокименко А. А.** Нахождение центра тяжести области работоспособности, заданной в матричной форме // Информатика и системы управления. 2006. № 1 (11). С. 174—181.
9. **Саушев А. В.** Метод и алгоритмы параметрического синтеза электротехнических систем по критерию запаса работоспособности // Информационные технологии. 2012. № 12. С. 24—29.
10. **Назаров Д. А.** Использование распределенных вычислений при построении области работоспособности // Информатика и системы управления. 2008. № 1 (15). С. 142—151.
11. **Schiff J. L.** Cellular automata: a discrete view of the world. A John Wiley & Sons inc, Publication. University of Auckland. 2008, 279 p.

**Ya. V. Katueva**, Senior Researcher, gloria@iacp.dvo.ru, **D. A. Nazarov**, Researcher, nazardim@iacp.dvo.ru  
Institute of Automation and Control Processes, Far Eastern Branch of Russian Academy of Sciences, Vladivostok

## The Methods of Parametric Synthesis Based on Grid Representation of a Region of Acceptability

*An application of parametric synthesis methods to grid approximation of a region of acceptability is discussed in this paper. An implementation of analytical method of mechanical and electrical analogies for discrete region representation is offered. The methods of objective function calculation based on inscribing convex symmetrical figures of maximal volume into a region are proposed for maximal performance capacity criterion. The method of narrowing areas is implemented for discrete narrowing of a region boundary.*

**Keywords:** quality, reliability, system performance, parametric synthesis, region of acceptability

### References

1. **Norenkov I. P., Manichev V. B.** *Osnovy teorii i proektirovaniya SAPR*: ucheb dlja vtuzov po spec / Moscow, Vyssh. shk., 1990, 335 p.
2. **Abramov O. V., Katueva Ja. V., Nazarov D. A.** Optimal'nyj parametricheskij sintez po kriteriju zapasa rabotosposobnosti, *Problemy upravlenija*, 2007, no. 6, pp. 64—69.
3. **Saushev A. V.** *Parametricheskij sintez jelektrotehnicheskikh ustrojstv i sistem*, SPb.: GUMRF imeni admirala S. O. Makarova, 2013, 315 p.
4. **Katueva Ja. V., Nazarov D. A.** Algoritmy analiza oblasti rabotosposobnosti, zadannoj v matrichnoj forme, *Informatika i sistemy upravlenija*, 2005, no. 2 (10), pp. 118—128.
5. **Saushev A. V.** Obласти работоспособности jelektrotehnicheskikh sistem, Sankt-Peterburg, Politehnika, 2013, 414 p.
6. **Nazarov D. A.** Ispol'zovanie oblastej rabotosposobnosti dlja optimal'nogo vybora nominalov parametrov, *Informatika i sistemy upravlenija*, 2011, no. 2 (28), pp. 59—69.
7. **Anop M. F., Katueva Ja. V., Nazarov D. A.** Ispol'zovanie oblasti kachestva v zadachah obespechenija parametricheskoj nadjzohnosti, *Informatika i sistemy upravlenija*, 2014, no. 2 (40), pp. 66—74.
8. **Katueva Ja. V., Evdokimenko A. A.** Nahozhdenie centra tjazhesti oblasti rabotosposobnosti, zadannoj v matrichnoj forme, *Informatika i sistemy upravlenija*, 2006, no. 1 (11), pp. 174—181.
9. **Saushev A. V.** Metod i algoritmy parametricheskoto sinteza jelektrotehnicheskikh sistem po kriteriju zapasa rabotosposobnosti, *Informacionnye tehnologii*, 2012, no. 12, pp. 24—29.
10. **Nazarov D. A.** Ispolzovanie raspredelennykh vychisleniy pri postroenii oblasti rabotosposobnosti, *Informatika i sistemy upravleniya*, 2008, no. 1 (15), pp. 142—151.
11. **Schiff J. L.** *Cellular automata: a discrete view of the world*. A John Wiley & Sons inc, Publication, University of Auckland, 2008, 279 p.

# МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 519.863

**И. Л. Каширина**<sup>1</sup>, д-р техн. наук, доц., kash.irina@mail.ru,

**Я. Е. Львович**<sup>2</sup>, д-р техн. наук, проф., office@vvt.ru,

**С. О. Сорокин**<sup>3</sup>, зам. директора департамента, sorokin-so@mon.gov.ru

<sup>1</sup> Воронежский государственный университет, Воронеж

<sup>2</sup> Воронежский институт высоких технологий, Воронеж

<sup>3</sup> Министерство образования и науки Российской Федерации, Москва

## Модели и численные методы оптимизации формирования эффективной сетевой системы с кластерной структурой

*Рассмотрены механизмы повышения эффективности сетевых систем с кластерной структурой, использующие интегральные оценки элементов системы и основанные на решении ресурсной (параметрической), структурно-ресурсной и комбинированной задач дискретной оптимизации. Для отыскания решения каждой из задач предложены соответствующие алгоритмы.*

**Ключевые слова:** мониторинг, сетевая система, интегральное оценивание, ресурсная оптимизация, структурная оптимизация

### Введение

Современные формы реализации систем образования, медицинского обслуживания, торговли, сервиса характеризуются принадлежностью к общему классу сетевых систем с кластерной структурой. Эффективность функционирования и развития объединенных в сеть объектов  $O_i$  ( $i = \overline{1, I}$ ) характеризуется набором мониторируемых показателей  $y_{ij}$  ( $i = \overline{1, I}, j = \overline{1, J}$ ), на основе которых формируется интегральная оценка объекта  $Y_i = F(y_{ij})$  ( $i = \overline{1, I}$ ), определенная на заданном числовом интервале  $[A, O]$  [1, 2].

Интегральная оценка  $Y_i$  используется для разбиения всей совокупности объектов  $O_i$  ( $i = \overline{1, I}$ ) на  $M$  кластеров. Каждый кластер характеризуется определенным диапазоном изменений интегральной оценки  $Y_i$  на числовом интервале  $[A, O]$ . Нумерация кластеров проводится по следующему принципу: в кластер с номером  $m = 1$  входят объекты-лидеры, значение интегральной оценки  $Y_i^1$  лучшего из которых максимально и близко  $O$ ; кластер с номером  $m = M$  составляют объекты-аутсайдеры, значение интегральной оценки  $Y_i^M$  худшего из которых минимально и близко к  $A$ .

После упорядочения объектов по значению интегральной оценки и их последующей кластерной структуризации создаются предпосылки для оптимизации эффективности системы в целом. Пред-

лагается рассмотреть оптимизационные модели и численные методы формирования оптимальной сетевой системы по двум составляющим эффективности: структурной и ресурсной. Повышение структурной эффективности достигается трансформацией множества объектов  $O_i$  ( $i = \overline{1, I}$ ) в другое множество  $O_{i_1}$  ( $i_1 = \overline{1, I_1}$ ), где  $I_1 < I$ , отвечающее заданным условиям, а повышение ресурсной эффективности — за счет такого распределения общего для системы ресурсного обеспечения, при котором осуществляется перемещение объектов из кластеров с меньшим значением интегральной оценки в кластеры с большим ее значением.

### Оптимизация структурной эффективности

Задача оптимизации структурной эффективности сетевой системы состоит в ее редукции в целях выравнивания эффективности деятельности объектов  $O_i$  ( $i = \overline{1, I}$ ) таким образом, чтобы повысить уровень минимальной оценки  $Y_i^M$ . Для решения этой задачи используется следующий интеграционный механизм: в кластерах нижнего уровня выделяется некоторое подмножество объектов  $O_t$  ( $t = \overline{1, T}$ ) с низкими интегральными оценками  $Y_t$  ( $t = \overline{1, T}$ ), которые могут быть поглощены объектами-лидерами с высоким уровнем интегральной оценки  $Y_l$  ( $l = \overline{1, L}$ ),  $L = T$ ,  $Y_l \gg Y_t$ . Новый объект характеризуется интегральной оценкой  $Y_{lt} = f(Y_l, Y_t) < Y_l$

( $l = \overline{1, L}, t = \overline{1, T}$ ). Как правило, формула получения интегральной оценки нового объекта имеет вид:  $Y_l = c_1 Y_l + c_2 Y_r$ ,  $c_1 + c_2 = 1$ ,  $c_1, c_2 \geq 0$ . Этот механизм трансформируется в оптимизационную постановку: необходимо обеспечить такое поглощение каждым из объектов  $O_l$  одного из объектов  $O_r$ , чтобы суммарное снижение уровня интегральной оценки объектов-лидеров было минимальным. Такой постановке соответствует оптимизационная модель с булевыми переменными [2]

$$x_{lt} = \begin{cases} 1, & \text{если объект } O_l \text{ поглощает объект } O_r, \\ 0, & \text{в противном случае,} \end{cases} \quad (1)$$

целевой функцией

$$\sum_{l=1}^L \sum_{t=1}^T (Y_l - Y_{lt} x_{lt}) \rightarrow \min \quad (2)$$

и ограничениями

$$\sum_{l=1}^L x_{lt} = 1, t = \overline{1, T}, \quad (3)$$

$$\sum_{t=1}^T x_{lt} = 1, l = \overline{1, L}, \quad (4)$$

$$x_{lt} \in \{0, 1\}, t = \overline{1, T}, l = \overline{1, L}. \quad (5)$$

Поскольку оценки  $Y_l$  в рамках данной модели являются константами, целевая функция может быть переписана в виде

$$\sum_{l=1}^L \sum_{t=1}^T Y_{lt} x_{lt} \rightarrow \max. \quad (6)$$

Задача (1)–(5) относится к известному классу задач о назначении, она может быть эффективно решена, например венгерским методом [3].

Для продолжения интеграционного процесса в рамках редуцированной сети с числом объектов  $I_1 = I - T$  выделяем новую группу объектов-лидеров  $O_{l_1}$  ( $l_1 = \overline{1, L_1}$ ) и формируем новое множество объектов  $O_{t_1}$  ( $t_1 = \overline{1, T_1}$ ),  $T_1 = L_1$  и задачу (1)–(5) решаем вновь с введением булевых переменных, аналогичных (1). При продолжении процедуры редукации сети возможен вариант поглощения объектом-лидером нескольких объектов с низким уровнем эффективности.

### Оптимизация ресурсной эффективности

Задача оптимизации ресурсной эффективности сетевой системы с кластерной структурой состоит в выборе двухэтапного механизма распределения ресурсного обеспечения  $R$ : на первом этапе между кластерами, а на втором этапе между объектами оптимизированной сети. Распределение ресурса  $R$  между кластерами  $R_m$  ( $m = \overline{1, M}$ ) целесообразно

осуществить по принципу обратных приоритетов [4] с учетом суммарной потребности объектов кластера и приоритета каждого кластера, установленного в системе, а внутри кластера разделить ресурс на поддерживающий  $R_1^m$  и развивающий  $R_2^m$ . Предполагается поддерживающий ресурс также распределить по принципу обратных приоритетов с учетом потребностей объекта и значения интегральной оценки эффективности. Для распределения второго вида ресурса используется оптимизационный подход, ориентированный на возможность повышения уровня эффективности объекта  $O_{t_m}$  ( $t_m = \overline{1, T_m}$ ,

$m = \overline{2, M}$ ) по  $j$ -му показателю, наиболее близкому к среднему значению этого показателя по кластеру  $m = 1$  объектов-лидеров. Для этого введем коэффициент близости значения  $j$ -го показателя объекта  $O_{t_m}$  ( $t_m = \overline{1, T_m}$ ,  $m = \overline{2, M}$ ) к значениям этого по-

казателя у объектов-лидеров:  $a_{t_m j} = \frac{y_{t_m j}}{\bar{y}_{1j}}$ , где  $\bar{y}_{1j}$  — среднее значение  $j$ -го показателя для объектов 1-го кластера, и булевы переменные

$$x_{t_m j} = \begin{cases} 1, & \text{если объекту } O_{t_m} \text{ выделяется разви-} \\ & \text{вающий ресурс для повышения эффек-} \\ & \text{тивности по } j\text{-му показателю,} \\ 0, & \text{в противном случае, } t_m = \overline{1, T_m}, j = \overline{1, J}. \end{cases}$$

Тогда оптимизационная модель для объектов  $m$ -го кластера имеет следующий вид:

$$\sum_{t_m=1}^{T_m} \sum_{j=1}^J a_{t_m j} x_{t_m j} \rightarrow \max; \quad (7)$$

$$\sum_{t_m=1}^{T_m} \sum_{j=1}^J r_{t_m j} x_{t_m j} \leq R_m^2; \quad (8)$$

$$\sum_{j=1}^J x_{t_m j} = 1, t_m = \overline{1, T_m}; \quad (9)$$

$$x_{t_m j} \in \{0, 1\}, t_m = \overline{1, T_m}, j = \overline{1, J}. \quad (10)$$

Целевая функция (7) определяет претендентов на развивающий ресурс  $R_m^2$ , максимально близких по показателям эффективности к объектам-лидерам, а ограничение (8) характеризует условие, при котором суммарная потребность в ресурсе объектов  $O_{t_m}$  для повышения уровня  $j$ -го показателя  $r_{t_m j}$  не должна превышать выделенного для  $m$ -го кластера ресурса  $R_m^2$ . Ограничение (9) означает, что каждому объекту дополнительный ресурс первонач-

чально выделяется для повышения только одного показателя.

Очевидно, что задача (7)–(10) разрешима только

в том случае, если  $\sum_{t_m=1}^{T_m} \min_{j=1, \dots, J} r_{t_m j} \leq R_m^2$ . Кроме того,

задача имеет смысл только в том случае, если

$R_m^2 < \sum_{t_m=1}^{T_m} \max_{j=1, \dots, J} r_{t_m j}$ , иначе ее решение тривиально:

$x_{t_m k} = 1$ , если  $k = \arg \max_{1 \leq j \leq J} a_{t_m j}$ ;  $x_{t_m j} = 0$  для  $j \neq k$ .

Отметим, что задача (7)–(10) близка по смыслу к общей постановке блочной задачи о рюкзаке (*Multiple-choice Knapsack Problem — МСКР*) [8, 9]. В этой задаче дано  $m$  классов  $N_1, N_2, \dots, N_m$  предметов, которые надо упаковать в рюкзак емкостью  $W$ . Каждый предмет  $j$  класса  $N_i$  имеет определенную ценность и массу. Задача заключается в том, чтобы выбрать ровно один предмет из каждого класса таким образом, что суммарная ценность предметов была максимальна и масса рюкзака не превышала  $W$ .

Блочная задача о рюкзаке является *NP*-сложной. Несмотря на довольно высокую практическую значимость *МСКР* (она возникает, например, при планировании инвестиций, проектировании информационных систем, оптимальном резервировании элементов сложных систем [6], планировании работ серверов и т. д.), число подходов к ее решению, описанных в литературе, остается сравнительно небольшим. Тем не менее существуют несколько точных методов для решения блочной задачи. Однако в случае большой размерности исходной постановки, точные методы решения *NP*-полных задач становятся практически неприменимыми.

В данной работе для решения задачи (7)–(10) предлагается приближенный метод, основанный на идее генетических алгоритмов.

Прежде всего отметим, что если в исходной постановке задачи (7)–(10) существуют номера  $l$  и  $k$  такие, что  $a_{t_m l} \leq a_{t_m k}$  и  $r_{t_m l} \geq r_{t_m k}$ , то вариант  $t_m l$  доминируется вариантом  $t_m k$ , а следовательно, переменная  $x_{t_m l}$  в оптимальном решении примет значение 0 (так как объект  $O_{t_m}$  при меньшем количестве ресурсов может сильнее повысить показатель  $k$ , чем показатель  $l$ ).

На рис. 1 изображены (в качестве примера) все возможные варианты повышения одного показателя эффективности для объекта  $O_{t_m}$  для всех  $j = \overline{1, J}$  (по оси абсцисс откладывается потребность в ресурсе  $r_{t_m j}$  для повышения уровня  $j$ -го показателя, по оси ординат — значение соответствующего повышения). Черным цветом выделены недомини-

руемые варианты, один из которых должен быть выбран в результате решения задачи.

Обозначим через  $N_{t_m}$  подмножество индексов показателей  $j$ , которым соответствуют недоминируемые варианты  $t_m j$  повышения уровня эффективности объекта  $O_{t_m}$ . Упорядочим элементы каждого множества  $N_{t_m}$  в порядке убывания  $a_{t_m j}$ . Например, если  $N_{t_m} = (3, 1, 5)$  — это значит что  $a_{t_m 3} \geq a_{t_m 1} \geq a_{t_m 5}$ . При этом элементы множеств  $N_{t_m}$  также автоматически станут упорядочены в порядке убывания  $r_{t_m j}$ , в силу наличия в  $N_{t_m}$  только недоминируемых вариантов, если  $a_{t_m l} \geq a_{t_m k}$ , то и  $r_{t_m l} \geq r_{t_m k}$ .

Кодирование данных для генетического алгоритма предлагается осуществлять следующим образом. Экземпляр популяции — это целочисленная строка длины  $T_m$ , в которой на месте  $t_m$  ( $t_m = \overline{1, T_m}$ ) стоит  $j \in N_{t_m}$ , если  $x_{t_m j} = 1$ . Таким образом, решение

$X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$  (фенотип) будет записано как строка

вида  $S = (1, 3, 4, 3)$  (генотип). Для оценки особей популяции по критерию приспособленности можно непосредственно использовать целевую функцию (7), вычисленную для соответствующего фенотипа. Для реализации процедур селекции и скрещивания можно использовать стандартные механизмы, а в процедуре мутации случайно выбранный  $k$ -й элемент строки заменить на другой случайный элемент из множества  $N_k$ .

Однако ввиду наличия ограничения (8) решения-потомки могут оказаться недопустимыми, поэтому необходима дополнительная процедура исправления недопустимых решений. Предположим,

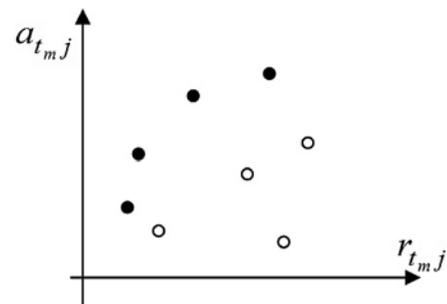


Рис. 1. Множество вариантов повышения одного из показателей эффективности для объекта  $O_{t_m}$

матрица  $X$ , соответствующая строке  $S$ , нарушает ограничение (8).

1. Определяем для каждого элемента  $s_{t_m}$  ( $t_m = \overline{1, T_m}$ ) строки  $S$  порядковый номер  $l$  в множестве  $N_{t_m}$  (например, если  $N_{t_m} = (3, 1, 5)$ , а в строке  $S$  на месте  $t_m$  стоит 5, то  $l = 3$ ).

2. В строке  $S$  заменяем некоторый элемент  $s_k$ , так, чтобы уменьшить требования к ресурсу, т. е. берем в множестве  $N_k$  вместо  $s_k$  следующий по порядку элемент.

Очевидно в качестве  $s_k$  нельзя выбирать элементы, которые имеют максимальный порядковый номер в множестве  $N_k$  (и, соответственно, минимальные требования к ресурсу для повышения  $k$ -го показателя). Поэтому предлагается для замены выбрать элемент  $s_k$  ( $k = \overline{1, T_m}$ ), которому соответствует минимальное отношение

$$d_k = \frac{a_{kl} - a_{k(l+1)}}{r_{kl} - r_{k(l+1)}}, \quad (11)$$

т. е. определяется замена, которая обеспечит минимальную потерю эффективности на единицу уменьшения требований к ресурсам.

Рассмотрим простой пример, иллюстрирующий предлагаемую процедуру исправления недопустимых решений. Пусть в задаче (7)–(10)  $T_m = J = 3$ ,

$R_m^2 = 10$ , при этом матрицы  $R$  и  $A$  имеют вид:

$$R = \begin{pmatrix} 4 & 3 & 5 \\ 6 & 2 & 1 \\ 5 & 3 & 2 \end{pmatrix}, A = \begin{pmatrix} 0,1 & 0,3 & 0,8 \\ 0,1 & 0,4 & 0,3 \\ 1 & 0,9 & 0,3 \end{pmatrix}. \text{ Определим мно-}$$

жества  $N_{t_m}$ ,  $t_m = \overline{1, 3}$ . На рис. 2 элементы этих множеств изображены графически и выделены черным цветом.

Упорядочив элементы этих множеств в порядке убывания  $a_{t_m j}$ , получаем:  $N_1 = \{3, 2\}$ ,  $N_2 = \{2, 3\}$ ,  $N_3 = \{1, 2, 3\}$ . Пусть на очередном шаге алгоритма

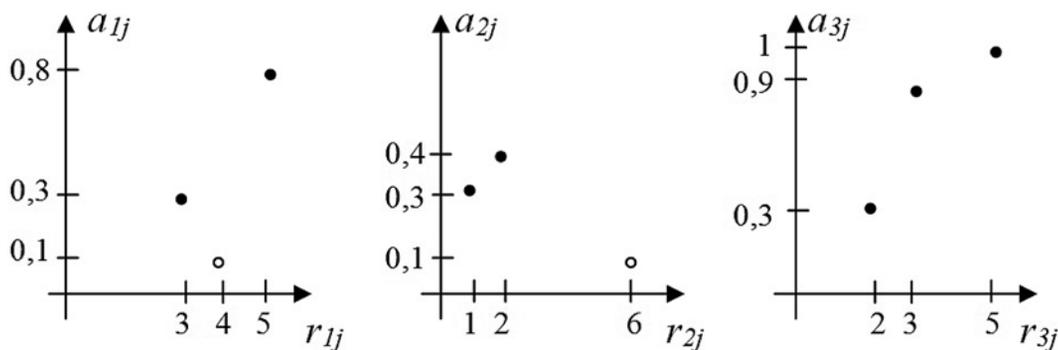


Рис. 2. Элементы множеств  $N_{t_m}$ ,  $t_m = \overline{1, 3}$

получена строка  $S = (3, 3, 1)$ , нарушающая ограничение (8):  $5 + 1 + 5 > 10$ . Определим подходящий элемент для замены. Очевидно это не может быть  $s_2$ , так как  $r_{23} = \min_{1 \leq j \leq 3} r_{2j}$ . Сравним  $s_1$  и  $s_3$ :

$$d_1 = \frac{0,8 - 0,3}{5 - 3} = 0,25, d_3 = \frac{1 - 0,9}{4 - 3} = 0,1. \text{ Таким образом, элемент } s_3 \text{ заменяется на следующее после него значение из } N_3, \text{ и строка } S \text{ принимает вид } S = (3, 3, 2).$$

Эта строка удовлетворяет ограничению (8) (в противном случае процедура корректировки решения осуществляется повторно).

### Оптимизация структурно-ресурсной эффективности

В целом ряде случаев целесообразно использовать комбинированный механизм оптимизации сетевой системы. С этой целью объекты  $O_{t_m}$   $m$ -го кластера

разбивают на две группы  $O_{t_m}^1$  ( $t_m^1 = \overline{1, T_m^1}$ ) и  $O_{t_m}^2$

( $t_m^2 = \overline{1, T_m^2}$ ), при этом  $T_m^1 + T_m^2 = T_m$ . В первую группу входят объекты, которые имеют достижения по отдельным показателям эффективности, близкие к объектам группы лидеров, и при выделении развивающего ресурса способны самостоятельно улучшить свою позицию в рейтинге. Во вторую группу включены объекты, которые самостоятельно не могут улучшить свое положение и подлежат поглощению объектами-лидерами. После разделения на группы для объектов первой группы используется оптимизационная модель (7)–(10), а второй группы — модель (2)–(5).

При разделении на группы надо учитывать следующие условия:

- ограниченность развивающего ресурса  $R_m^2$ , что определяет ограниченность подмножества объектов первой группы;
- объекты, включаемые в первую группу, должны иметь перспективу роста по каждому  $j$ -му показателю.

Если ввести бинарную оценку выполнения второго условия каждым объектом

$$c_{t_m j} = \begin{cases} 1, & \text{если объект } O_{t_m} \text{ имеет перспективу} \\ & \text{приближения по } j\text{-му показателю к объ-} \\ & \text{ектам-лидерам,} \\ 0, & \text{в противном случае, } t_m = \overline{1, T_m}, j = \overline{1, J} \end{cases}$$

и булевы переменные

$$x_{t_m} = \begin{cases} 1, & \text{если объект } O_{t_m} \text{ включается в первую} \\ & \text{группу,} \\ 0, & \text{в противном случае, } t_m = \overline{1, T_m}, \end{cases}$$

то оптимизационная модель, обеспечивающая выполнение обоих условий, имеет следующий вид:

$$\sum_{t_m=1}^{T_m} x_{t_m} \rightarrow \min; \quad (12)$$

$$\sum_{t_m=1}^{T_m} c_{t_m j} x_{t_m} \geq 1, j = \overline{1, J}; \quad (13)$$

$$x_{t_m} \in \{0, 1\}, t_m = \overline{1, T_m}, \quad (14)$$

где целевая функция (12) определяет выполнение первого условия, а ограничение (13) — выполнение второго условия.

Задача (12)—(14) относится к классу задач о минимальном покрытии и для ее решения предлагается использовать оригинальный подход, предложенный в работе [5]. На основании этого подхода задача дискретной оптимизации (12)—(14) может быть эквивалентным образом переписана в виде задачи непрерывной минимизации полинома, заданного на единичном гиперкубе:

$$-\sum_{t_m=1}^{T_m} q_{t_m} + T_m \sum_{j=1}^J \prod_{t_m: c_{t_m j}=1} q_{t_m} \rightarrow \min_{0 \leq q_{t_m} \leq 1}, \quad (15)$$

при этом координаты вектора  $x_{t_m}$ , оптимального в задаче (12)—(14), связаны с координатами оптимального вектора  $q_{t_m}$  соотношением:  $x_{t_m} = 1 - q_{t_m}$ , т. е. решением непрерывной задачи (15) будет вектор с булевыми координатами. Такой переход осуществляется путем рандомизации переменных и рассмотрения вероятностной постановки исходной задачи (12)—(14). При этом для решения задачи (15) может быть использован, например, градиентный алгоритм с проекцией на единичный гиперкуб. Надо отметить, что оптимизируемая градиентным алгоритмом

функция в задаче (15) не имеет локальных экстремумов внутри единичного гиперкуба [5], но они могут присутствовать в граничных точках. Поэтому в целях отыскания глобального экстремума для каждой решаемой задачи градиентный алгоритм рекомендуется перезапускать несколько раз, каждый раз с новой случайной начальной точки.

## Заключение

По результатам вычислительного эксперимента можно отметить, что при решении тестовых задач предложенными методами обеспечивается достаточно высокая точность. Так, для тестовых задач (7)—(10), в которых было известно оптимальное решение, отношение найденного генетическим алгоритмом приближенного значения целевой функции к оптимальному, в среднем, превышало 0,9. Градиентный алгоритм решения задачи (15) на сильно разреженных матрицах покрытия (содержащих более 80 % нулевых элементов) работает значительно быстрее метода ветвей и границ, т. е. традиционного алгоритма решения задач о минимальном покрытии.

Таким образом, совокупность предлагаемых оптимизационных моделей и соответствующего алгоритмического обеспечения позволяет осуществить формирование эффективной сетевой системы с кластерной структурой.

## Список литературы

1. Львович Я. Е. Многоальтернативная оптимизация: теория и приложения. Воронеж: Кварта, 2006. 428 с.
2. Каширина И. Л., Львович Я. Е., Сорокин С. О. Интегральное оценивание эффективности сетевых систем с кластерной структурой // Экономика и менеджмент систем управления. 2015. Т. 15. № 1.3. С. 330—337.
3. Батищев Д. И., Львович Я. Е., Фролов В. Н. Оптимизация в САПР. М.: Высшая школа, 1997. 416 с.
4. Бурков В. Н., Новиков Д. А. Теория активных систем: состояние и перспективы. М.: Синтег, 1999. 128 с.
5. Львович Я. Е., Каширина И. Л., Чернышова Г. Д. Оптимизация проектных решений на основе эквивалентных преобразований задачи о минимальном покрытии // Информационные технологии. 1999. № 4. С. 2—6.
6. Львович Я. Е., Каширина И. Л., Тузиков А. А. Генетический алгоритм решения многокритериальной задачи повышения надежности резервирования // Информационные технологии. 2012. № 6. С. 56—60.
7. Каширина И. Л., Львович Я. Е., Сорокин С. О. Ресурсная оптимизация эффективности сетевых систем с кластерной структурой // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2015. № 1. С. 39—43.
8. Левин М. Ш., Сафонов А. В. Эвристический алгоритм для многокритериальной блочной задачи о рюкзаке // Искусственный интеллект и принятие решений. 2009. № 4. С. 53—64.
9. Dyer M. E., Kayal N., Walker J. A Branch and Bound Algorithm for Solving the Multiple-Choice Knapsack Problem // Journal of Computational and Applied Mathematics. 1984. N. 11. P. 231—249.

**I. L. Kashirina**, Associate Professor, Voronezh state University, e-mail: kash.irina@mail.ru,  
**Ja. E. Lvovich**, PhD, Professor, President, Voronezh Institute of High Technologies, e-mail: office@vivt.ru,  
**S. O. Sorokin**, Deputy Director of the Department, Ministry of Education and Science  
of the Russian Federation, e-mail: sorokin-so@mon.gov.ru

## Models and Numerical Methods Optimization of Formation Effective Network System Cluster Structure

*The article deals with mechanisms to improve the efficiency of network systems with cluster structure using integrated assessment system elements and based on the solution of the resource (parametric) and structural resources and the combination of discrete optimization problems.*

*The problem of structural and resource optimization consists in the reduction of the network structure in order to equalize the efficiency of the facilities so as to increase the level of the minimum of the integral evaluation. The mathematical model of the problem is close to setting a block of the knapsack problem. To solve it, a genetic algorithm is proposed.*

*The problem of resource optimization is to select a two-step mechanism for the allocation of resource provision: in the first stage, between clusters, and then between objects optimized network.*

*For the distribution of resource use optimization-based approach to the possibility of raising the level of asset performance indicators, the closest to the values of the relevant indicators of leaders. Mathematical model of the problem belongs to the class of problems of appointment. To solve it, can be used Hungarian method.*

*Combined optimization mechanism network system is to find the optimum separation of the cluster objects with the lowest values of integral evaluation into two groups. The first group includes objects that the allocation of developmental resources, the ability to improve their own position in the ranking. The second group included subjects who alone can not improve his position and to be absorbed by an object-leaders. The mathematical model of the problem belongs to a class of problems on the minimal covering. To solve it, the original algorithm is proposed.*

*Thus, the totality of the proposed optimization models and algorithms will allow to carry out the decision support for the cluster management of network system.*

**Keywords:** monitoring, network system, integrated estimation, resource optimization, structural optimization

### References

1. **Lvovich Ja. E.** *Mногоальтернативная оптимизация: теория и приложения* (Multialternative Optimization: Theory and Applications), Voronezh, Kvarta, 2006, 428 p. (in Russian).
2. **Kashirina I. L., Lvovich Ja. E., Sorokin S. O.** *Jekonomika i menedzhment sistem upravlenija*, 2015, vol. 15, no. 1.3, pp. 330–337 (in Russian).
3. **Batishhev D. I., Lvovich Ja. E., Frolov V. N.** *Optimizacija v SAPR* (Optimization in CAD), Moscow, Vysshaja shkola, 1997, 416 p. (in Russian).
4. **Burkov V. N., Novikov D. A.** *Teorija aktivnyh sistem: sostojanie i perspektivy* (Theory of active systems: state and prospects), Moscow, Sinteg, 1999, 128 p. (in Russian).
5. **Lvovich Ja. E., Kashirina I. L., Chernyshova G. D.** *Informacionnye tehnologii*, 1999, no. 4, pp. 2–6. (in Russian).
6. **Lvovich Ja. E., Kashirina I. L., Tuzikov A. A.** *Informacionnye tehnologii*, 2012, no. 6, pp. 56–60. (in Russian).
7. **Kashirina I. L., Lvovich Ja. E., Sorokin S. O.** *Vestnik Voronezhskogo gosudarstvennogo universiteta. Serija: Sistemnyj analiz i informacionnye tehnologii*, 2015, no. 1, pp. 39–43. (in Russian).
8. **Levin M. Sh., Safonov A. V.** *Iskusstvennyj intellekt i prinjatje reshenij*, 2009, no. 4, pp. 53–64. (in Russian).
9. **Dyer M. E., Kayal N., Walker J.** A Branch and Bound Algorithm for Solving the Multiple-Choice Knapsack Problem, *Journal of Computational and Applied Mathematics*, 1984, no. 11, pp. 231–249.

УДК 656.073

**Е. М. Бронштейн**, д-р физ.-мат. наук, проф., e-mail: bro-efim@yandex.ru,  
**И. Е. Копылов**, аспирант, e-mail: kopilya@gmail.com,  
Уфимский государственный авиационный технический университет

## Оптимизационная задача транспортной логистики с простыми маршрутами передвижения

*Рассмотрена задача построения оптимального маршрута доставки однородного груза от семейства производителей семейству потребителей при расположении пунктов на линейной или кольцевой дорогах одним транспортным средством с ограниченной вместимостью. Проведен сравнительный анализ трех эвристических алгоритмов.*

**Ключевые слова:** транспортная логистика, оптимизация, эвристические методы, метод ветвей и границ, жадный алгоритм

## Введение

В связи с ростом городов усложняется задача доставки грузов из пунктов производства в пункты потребления. Для уменьшения затрат на транспортировку груза необходимо выбрать наилучший в том или ином смысле маршрут движения транспортных средств (ТС).

На сегодняшний день сформулировано множество подобных задач, в которых учитываются различные ограничения, в том числе на вместимость ТС, на число ТС, на время доставки грузов, на расположение пунктов и т. д. Классификация оптимизационных задач транспортной логистики приведена в работе [1].

Сети дорог могут быть различными. Простейшим случаем является линейное (кольцевое) расположение пунктов. Такое расположение пунктов особенно актуально для России, где сеть дорог весьма бедная.

Подобные задачи рассматривались и ранее. Например, в [2] была рассмотрена задача доставки грузов по кольцевой дороге при условии, что суммарный объем наличного товара и общий объем требуемого товара не превышают вместимости ТС.

Целью работы является формализация задачи и сравнительный анализ некоторых эвристических алгоритмов.

## Постановка задачи

Пусть есть несколько пунктов производства и потребления. Все пункты находятся на одной линии. Вариант — линия замкнута (пункты расположены на кольцевой дороге). Одному ТС необходимо развести однородный груз из пунктов производства в пункты потребления.

ТС начинает движение из депо. После развозки грузов ТС должно в депо вернуться. Вместимость ТС ограничена. Предполагается, что операции загрузки и выгрузки груза являются весьма трудоемкими, поэтому каждый пункт обслуживается только один раз.

ТС может забирать груз из пункта производства, если его вместимости достаточно для загрузки имеющегося груза (ТС не догружено на соответствующий объем груза). Естественно, выгрузку в пункте потребления можно проводить, только если в ТС груза достаточно. Хотя каждый пункт обслуживается ТС один раз, проезжать любой пункт можно многократно.

**Математическая модель задачи.** Есть  $n$  пунктов, которые находятся на одной линии (частный случай — пункты находятся на кольцевой дороге, условия, которые описывают данный случай, будем заключать в квадратные скобки). Считаем депо нулевым пунктом. По окончании доставки грузов ТС должно вернуться в депо.

Пусть  $t(i)$  — координата  $i$ -го пункта. Полагаем  $t(0) = 0$ ,  $t(i) \neq 0$  при  $i \neq 0$ ,  $[t(i) \geq 0]$ .

Пусть  $c_{ij}$  — расстояние между  $i$ -м и  $j$ -м пунктами. Для линейного расположения пунктов  $c_{ij} = |t(i) - t(j)|$ , [для кольцевого  $c_{ij} = \min\{|t(i) - t(j)|, L - |t(i) - t(j)|\}$ , где  $L$  — длина кольцевой дороги].

Масса груза (или потребность в грузе) в  $i$ -м пункте обозначим  $a(i)$ ,  $a(i) > 0$ , если  $i$ -й пункт — пункт производства,  $a(i) < 0$ , если  $i$ -й пункт — пункт потребления.

Считаем, что задача сбалансирована  $\left(\sum_{i=0}^n a(i) = 0\right)$ .

Пусть  $Q$  — вместимость ТС.

Необходимо найти последовательность номеров обслуживания пунктов в порядке их прохождения  $u(i)$  ( $i = 0, \dots, n + 1$ ).

Должны выполняться следующие условия:

- $u(0) = u(n + 1) = 0$

(движение начинается и заканчивается в депо);

- $u(i) \neq u(j)$  при  $i \neq j$ , ( $i, j = 0, \dots, n$ )

(подлежат обслуживанию все пункты по одному разу);

- $0 \leq \sum_{i=0}^k a[u(i)] \leq Q$  ( $k = 0, \dots, n - 1$ ), если  $a(0) \geq 0$ ,

- $0 \leq \sum_{i=0}^k a[u(i)] \leq Q$  ( $k = 1, \dots, n$ ), если  $a(0) < 0$

(ограничение на вместимость);

- $\sum_{i=0}^{n-1} c_{u(i), u(i+1)} + c_{u(n), 0} \rightarrow \min$

(длина маршрута должна быть минимальной).

В общем случае задача может не иметь решения. Для более общей задачи в работе [3] доказано, что при выполнении условия  $Q \geq 2 \max_i |a_i|$  задача имеет решение.

Далее полагаем, что  $Q = 2 \max_i |a_i|$ . В [3] доказано, что в этом случае любой допустимый маршрут, содержащий не все пункты, может быть продолжен.

## Методы решения задачи

Для решения задачи применялись три эвристических метода.

**Усеченный метод ветвей и границ.** На каждой итерации для каждой ветви решения создаются две новые ветви: ТС едет в ту или иную сторону. Если есть возможность в новом пункте загрузить или выгрузить груз, то осуществляется загрузка/разгрузка.

Устанавливается верхняя граница — длина некоторого маршрута ТС. Значение границы определяется из решения задачи другим эвристическим методом (жадным алгоритмом по расстоянию). Если в какой-либо ветви решения длина маршрута превышает принятую границу, данная ветвь исключается из рассмотрения.

Применяются некоторые условия отсечения ветвей. Первое условие для линейной и кольцевой

дорог: если ТС доезжает до пункта и, не совершив загрузку/выгрузку груза, едет в обратном направлении, то ветвь исключается из рассмотрения. Второе условие для линейной дороги: если ТС движется в направлении, в котором все пункты уже обслужены, соответствующая ветвь не рассматривается.

**Жадный алгоритм по расстоянию.** На каждом шаге алгоритма ищутся ближайшие пункты с двух сторон от текущего пункта, в которых ТС может забрать или оставить груз. ТС едет в ближайший из этих двух пунктов.

**Жадный алгоритм по маршруту.** ТС начинает движение в одном из двух направлений. По мере движения в каждом посещаемом пункте, если есть возможность, ТС проводит загрузку или выгрузку груза. При этом при посещении каждого пункта проверяется целесообразность движения далее. Если дальнейшее движение нецелесообразно, то ТС меняет направление движения. Условия смены направления движения:

- во все последующие пункты по направлению движения груз доставлен;
- в последующих пунктах ТС не сможет загрузить/выгрузить груз.

Сравниваются длины завершённых маршрутов для обоих направлений движения в начальный момент и выбирается меньшая.

Для случая кольцевого расположения пунктов этот алгоритм не рассматривается. При движении по кольцевой дороге ТС никогда не будет менять направление движения и развезет все грузы максимально за  $n$  циклов. Эксперименты показали, что данный алгоритм, как и следовало ожидать, менее эффективен, нежели другие, рассмотренные в статье.

Если изменить алгоритм построения маршрута по кольцевой дороге таким образом, что ТС движется до ближайшего пункта, допустимого для обслуживания, то алгоритм не отличается от жадного алгоритма по маршруту.

### Вычислительный эксперимент

Описанные алгоритмы были реализованы с помощью языка C#. Эксперимент проводился на компьютере с процессором Intel Core i7-3612QM CPU 2.10 GHz, оперативной памятью 8 GB.

Решалось по 30 случайно сгенерированных примеров для 10, 20, 30, 40, 50 пунктов. Задачи с большим числом пунктов решить задачей методом ветвей и границ не удалось из-за переполнения оперативной памяти. В связи с этим для 60, 70, 80, 90, 100 пунктов проводили только сравнительный анализ жадных алгоритмов на линейной дороге.

**Генерация примеров.** Для выбранного числа пунктов массы грузов во всех пунктах кроме нулевого, генерировались случайно в диапазоне от  $-20$  до  $+20$ . Масса груза в депо определялась из балансового соотношения.

Координата  $t(i)$  каждого пункта генерировалась случайным образом в диапазоне  $[-2n, 2n]$  для линейного расположения пунктов и в диапазоне  $[0, 4n]$  для кольцевого, где  $n$  — число пунктов. Длина кольцевой дороги принималась равной максимальной из координат пунктов с добавлением случайной величины из промежутка  $[0, 20]$ .

**Результаты вычислительного эксперимента.** Результаты эксперимента представлены в табл. 1–3. Метод ветвей и границ в среднем порождал относительно более короткий маршрут для рассмотренных задач, поэтому результаты по данному методу были приняты за эталон. Данные в таблицах — средние значения (среднее время работы программы, среднее отклонение результатов целевых функций от эталона в процентах).

Усеченный метод ветвей и границ не гарантирует получения оптимального решения — в некоторых случаях другие исследованные алгоритмы показывали лучший результат.

Всего было проведено по 150 экспериментов для линейной и кольцевой дорог при числе пунктов 10, 20, 30, 40, 50. Среди проведенных экспериментов в 15 случаях на линейной дороге и в одном случае на кольцевой жадный алгоритм по расстоянию порождал более короткий маршрут, чем усеченный метод ветвей и границ. Жадный алгоритм по маршруту порождал результат, лучший относительно усеченного метода ветвей и границ 17 раз на линейной дороге.

Результаты эксперимента на линейной дороге представлены в табл. 1.

Результаты эксперимента на кольцевой дороге представлены в табл. 2.

Для числа пунктов 60, 70, 80, 90, 100 использовали жадный алгоритм по маршруту и жадный алгоритм по расстоянию. Жадный алгоритм по расстоянию в среднем давал лучший результат, поэтому он был выбран за эталон (табл. 3).

Таблица 1

Число пунктов	Параметр	Ветви	Жадный по расстоянию	Жадный по маршруту
10	Время, с Отклонение результатов, %	0,01 0	0,0001 9,9	0,0001 4,8
20	Время, с Отклонение результатов, %	0,46 0	0,0001 12	0,0001 12,6
30	Время, с Отклонение результатов, %	10,98 0	0,0001 20,2	0,0001 21,3
40	Время, с Отклонение результатов, %	338,6 0	0,0001 20,6	0,0001 22,7
50	Время, с Отклонение результатов, %	14 618,7 0	0,0002 31,7	0,001 33,3

Таблица 2

Число пунктов	Параметр	Ветви	Жадный по расстоянию
10	Время, с	0,01	0,0001
	Отклонение результатов, %	0	12,4
20	Время, с	0,18	0,0002
	Отклонение результатов, %	0	29,5
30	Время, с	6,17	0,0003
	Отклонение результатов, %	0	37,6
40	Время, с	513,83	0,0003
	Отклонение результатов, %	0	37,8
50	Время, с	3910,93	0,0003
	Отклонение результатов, %	0	45,7

Таблица 3

Число пунктов	Параметр	Жадный по расстоянию	Жадный по маршруту
60	Время, с	0,0002	0,001
	Отклонение результатов, %	0	2,6
70	Время, с	0,0002	0,001
	Отклонение результатов, %	0	7,9
80	Время, с	0,0002	0,001
	Отклонение результатов, %	0	6,6
90	Время, с	0,0002	0,001
	Отклонение результатов, %	0	11,6
100	Время, с	0,0003	0,001
	Отклонение результатов, %	0	3,9

Из таблицы следует, что при применении жадного алгоритма по маршруту длина полученного маршрута в среднем на 6,5 % больше длины маршрута, полученного жадным алгоритмом по расстоянию.

### Заключение

В работе рассмотрена задача построения оптимального маршрута доставки однородного груза одним ТС с ограниченной вместимостью между  $n$  пунктами производства и потребления при условии расположения пунктов на линейной и кольцевой дорогах.

Проведен сравнительный анализ эффективности применения трех эвристических методов: усеченного метода ветвей и границ, жадного алгоритма по расстоянию и жадного алгоритма по маршруту.

Метод ветвей и границ в среднем дает лучшие результаты, но не применим для задач больших размерностей и существенно уступает другим алгоритмам по времени работы.

При числах пунктов, равных 60, 70, 80, 90, 100, жадный алгоритм по расстоянию работает быстрее и дает лучшие результаты нежели жадный алгоритм по маршруту.

*Работа выполнена при поддержке РФФИ (проект 01-13-00005).*

### Список литературы

1. **Бронштейн Е. М., Заико Т. А.** Детерминированные оптимизационные задачи транспортной логистики // Автоматика и телемеханика. 2010. № 10. С. 133–147.
2. **Gendreau M., Laporte G., Vigo D.** Heuristics for the traveling salesman problem with pickup and delivery // Computers & Operations Research. 1999. N. 26. P. 699–714.
3. **Бронштейн Е. М., Гиндуллин Р. В.** Об одном классе задач маршрутизации // Математическое моделирование. 2011. Т. 23, № 6. С. 122–123.

**E. M. Bronshtein**, Professor, e-mail: bro-efim@yandex.ru,

**I. E. Kopylov**, Postgraduate Student, e-mail: kopilya@gmail.com, Ufa State Aviation Technical University

## Vehicle Routing Problem with Simple Movement Route

*We consider the problem of constructing an optimal route of homogeneous cargo delivery from a family of producers to a family of consumers with the location of points on the line or the ring road by a vehicle with limited capacity. We give a comparative analysis of the three heuristic algorithms: truncated branch and bound method, greedy algorithm by distance and greedy algorithm by route. Branch and bound method gives better results, but is not suitable for large-scale problems when the number of points is more than 50. In this case, the branch and bound algorithms significantly inferior to the other in the counting time. When the number of points equal to 60, 70, 80, 90, 100 greedy algorithm for distance is faster and gives better results than the greedy algorithm by route.*

**Keywords:** transportation logistics, optimization, heuristic methods, branch and bound method, greedy algorithm

### References

1. **Bronshtein E. M., Zaiko T. A.** Deterministic optimization problems of transport logistics. *Automation and Remote Control*, 2010, no. 10, pp. 133–147.

2. **Gendreau M., Laporte G., Vigo D.** Heuristics for the traveling salesman problem with pickup and delivery. *Computers & Operations Research*, 1999, vol. 26, no. 7, pp. 699–714.

3. **Bronshtein E. M., Gindullin R. V.** About one class of routing problems. *Mathematical Modelling*, 2011, vol. 23, no. 6, pp. 122–123.

А. А. Моисеев, канд. техн. наук, ст. науч. сотр., e-mail: slow.coach@yandex.ru  
ГосНИИ химмотологии

## Моделирование химической стабильности жидкости методом теории подобия

*Проведено моделирование химической стабильности жидкости методом теории подобия. Целью моделирования являлось определение параметров поверочного эксперимента по оценке химической стабильности жидкости в гидростатическом случае. Оно осуществлялось путем построения соответствующих критериев: геометрического подобия экспериментального и натурального сосудов, термодинамического подобия температур и давлений, а также кинетического подобия протекания химической трансформации в натуральных и экспериментальных условиях. Полученные критерии позволяют оценить параметры поверочного эксперимента, обеспечивающие подобие реальным условиям химической трансформации жидкости, а также длительность указанного эксперимента.*

**Ключевые слова:** критерий подобия, химическая стабильность, химическая трансформация, геометрическое подобие, термодинамическое подобие, кинетическое подобие, температура, гидростатическое давление, поверочный эксперимент, уравнение Аррениуса

Теория подобия традиционно используется в исследованиях физических процессов в ситуации недостаточной проработанности теории последних. Наиболее широко она используется в механике [1], гидро- и термодинамике [2, 3], однако находит применение и в других областях, например, в физической химии [4, 5]. С точки зрения теории подобия моделирование физических процессов сводится к построению их безразмерных инвариантов — критериев подобия, а также соотношений между ними. Как правило, указанные инварианты формируются либо путем нормализации точных уравнений физических процессов [6], либо по результатам априорного анализа параметров, влияющих на процесс, с использованием  $\pi$ -теоремы [7, 8].

Цель исследования — определение параметров поверочного эксперимента по оценке химической стабильности жидкости в гидростатическом случае. Актуальность сформулированной цели обусловлена тем, что эксперимент предполагается провести при повышенной температуре для сокращения срока испытаний. Поэтому в число оцениваемых параметров эксперимента была включена его длительность.

### Синтез критериев подобия

Для достижения поставленной цели были построены критерии подобия, существенные для описания химической трансформации жидкости. При этом считалось, что в гидростатическом случае физико-химические процессы определяются следующим рядом параметров:

$$\rho = \frac{M}{L^3} \text{ — плотность жидкости;}$$

$$\mu = \frac{M}{N} \text{ — молярная масса жидкости;}$$

$$c = \frac{L^2}{\Theta T^2} \text{ — теплоемкость жидкости;}$$

$$E_0 = \frac{ML^2}{NT^2} \text{ — энергия активации реакции транс-}$$

формации;

$$V = L^3 \text{ — объем сосуда;}$$

$S = L^2$  — площадь внутренней поверхности сосуда;

$$P = \frac{M}{LT^2} \text{ — гидростатическое давление жидкости;}$$

$$T = \Theta \text{ — температура жидкости.}$$

В соответствии с выводами работы [7] только пять из этих параметров являются независимыми, а остальные определяются через них. Выбор независимых параметров базировался на алгебраическом доказательстве их независимости, а для построения зависимых критериев использована методика, представленная в работе [9].

В качестве независимых выберем параметры  $(\rho, \mu, c, E_0, V)$ , а в качестве зависимых — параметры  $(S, P, T)$ . Исходя из степенного представления  $\rho^\alpha \mu^\beta c^\gamma E_0^\delta V^\epsilon$ , получаем характеристический детерминант для независимых параметров [3]:

$$\det = \begin{vmatrix} M & 1 & 1 & 0 & 1 & 0 \\ L & -3 & 0 & 2 & 2 & 3 \\ T & 0 & 0 & -2 & -2 & 0 \\ \Theta & 0 & 0 & -1 & 0 & 0 \\ N & 0 & -1 & 0 & -1 & 0 \end{vmatrix} = -6 \neq 0.$$

Таким образом, выбранные параметры являются независимыми.

Система уравнений для зависимого параметра  $S$  имеет вид:

$$\begin{cases} \alpha + \beta + 0 + \delta + 0 = 0, \\ -3\alpha + 0 + 2\gamma + 2\delta + 3\varepsilon = 2, \\ 0 + 0 - 2\gamma - 2\delta + 0 = 0, \\ 0 + 0 - \gamma + 0 + 0 = 0, \\ 0 - \beta + 0 - \delta + 0 = 0. \end{cases}$$

Решая ее, получаем  $\alpha = 0$ ,  $\beta = 0$ ,  $\gamma = 0$ ,  $\delta = 0$ ,  $\varepsilon = \frac{2}{3}$ . Отсюда находим для характеристической [6]

площади поверхности  $S = V^{2/3}$ . Критерий при этом имеет вид  $\pi_1 = \frac{S}{V^{2/3}}$  и выражает отношение **геометрического подобия**.

Система уравнений для параметра  $P$  имеет вид:

$$\begin{cases} \alpha + \beta + 0 + \delta + 0 = 1, \\ -3\alpha + 0 + 2\gamma + 2\delta + 3\varepsilon = -1, \\ 0 + 0 - 2\gamma - 2\delta + 0 = -2, \\ 0 + 0 - \gamma + 0 + 0 = 0, \\ 0 - \beta + 0 - \delta + 0 = 0. \end{cases}$$

Решив ее, получим  $\alpha = 1$ ,  $\beta = -1$ ,  $\gamma = 0$ ,  $\delta = 1$ ,  $\varepsilon = 0$ . Отсюда находим характеристическое давление

$P = \frac{\rho E_0}{\mu}$ . Критерий при этом имеет вид  $\pi_2 = \frac{\mu P}{\rho E_0}$

и выражает условие подобия гидростатических давлений. Оно удовлетворяется либо выравниванием этих давлений, либо путем подбора альтернативной жидкости.

Система уравнений для  $T$  имеет вид:

$$\begin{cases} \alpha + \beta + 0 + \delta + 0 = 0, \\ -3\alpha + 0 + 2\gamma + 2\delta + 3\varepsilon = 0, \\ 0 + 0 - 2\gamma - 2\delta + 0 = 0, \\ 0 + 0 - \gamma + 0 + 0 = 1, \\ 0 - \beta + 0 - \delta + 0 = 0. \end{cases}$$

Решив ее, получим  $\alpha = 0$ ,  $\beta = -1$ ,  $\gamma = -1$ ,  $\delta = 1$ ,  $\varepsilon = 0$ . Отсюда находим характеристическую температуру

$T = \frac{E_0}{c\mu}$ . Критерий при этом имеет вид

$\pi_3 = \frac{c\mu T}{E_0}$  и выражает условие подобия температур.

Оно удовлетворяется либо выравниванием этих температур, либо путем подбора альтернативной жидкости.

Возможна ситуация, когда для ускорения процесса анализа в эксперименте используется повышенная температура. В этом случае непосредственное применение критериев  $\pi_2$ ,  $\pi_3$  нецелесообразно. Можно, однако, использовать их свертку следующего вида:

$$\pi_4 = \frac{\pi_2}{\pi_3} = \frac{P}{c\rho T}.$$

Этот критерий выражает условие **термодинамического подобия** — пропорциональности давлений и температур в моделируемом и моделирующем процессах. При произвольном значении температурного приращения  $\Delta T$  это условие выполняется для идеальных газов в изохорическом процессе [10]. Для жидкостей указанный критерий представляет собой условие выбора температурного приращения.

Пусть температура в моделируемом процессе равна  $T$ , а гидростатическое давление  $P = \rho g^3 \sqrt[3]{V}$ , где  $g$  — ускорение свободного падения. Предположим также, что экспериментальная колба объемом  $V_e$  наполнена той же жидкостью и закупорена. Тогда при температуре  $T + \Delta T$  давление в ней составит  $P_e = \rho g^3 \sqrt[3]{V_e} + E\alpha\Delta T$ , где  $E$  — модуль сжатия, а  $\alpha$  — коэффициент температурного расширения. Условие подобия в этом случае принимает вид:

$$\frac{\rho g^3 \sqrt[3]{V}}{T} = \frac{\rho g^3 \sqrt[3]{V_e} + E\alpha\Delta T}{T + \Delta T}.$$

Разрешая его относительно  $\Delta T$ , находим

$$\Delta T = T \frac{\rho g^3 (\sqrt[3]{V} - \sqrt[3]{V_e})}{E\alpha T - \rho g^3 \sqrt[3]{V}}.$$

Для слабосжимаемых жидкостей можно считать, что  $E\alpha T \gg \rho g^3 \sqrt[3]{V}$  [11]. Пренебрежем также  $\sqrt[3]{V_e}$  в сравнении с  $\sqrt[3]{V}$ , что дает погрешность в несколь-

ко процентов. Тогда  $\Delta T \approx \frac{\rho g^3 \sqrt[3]{V}}{E\alpha} \ll T$ . Это означает,

что критерий допускает лишь незначительное повышение температуры жидкости в наполненной и закупоренной колбе.

В случае если значение температурного приращения задано из каких-то дополнительных соображений, колба перед закупоркой при температуре  $T$  должна быть заполнена лишь частично, т. е. объем жидкости в ней  $V' < V_e$ . Заполнение колбы должно

происходить при температуре  $T + \Delta T$  за счет температурного расширения. Тогда

$$\begin{cases} \frac{\Delta V}{V'} = \alpha \Delta T, \\ V' + \Delta V = V_e. \end{cases}$$

Отсюда приближенно находим первоначальный объем жидкости в колбе:

$$V' = \frac{V_e}{1 + \alpha \Delta T}.$$

В рамках этого рассмотрения не учитывалась малая поправка приращения  $\Delta T'$ , требуемая для создания необходимого гидростатического давления. Эта поправка составляет

$$\Delta T' = \frac{\rho g (\sqrt[3]{V} - \sqrt[3]{V_e})}{E \alpha},$$

а температурное приращение в выражении для первоначального объема трансформируется в  $\Delta T - \Delta T'$ .

Оценим теперь необходимое время эксперимента. Будем исходить из того, что степень химической трансформации жидкости не должна превышать  $\eta$ . Учитывая, что скорость реакции определяется уравнением Аррениуса [12], находим выражение для характерного в смысле [6] времени трансформации

$$t = \frac{\mu \rho V \eta}{k_0 \exp\left(-\frac{E_0}{RT}\right)} = \frac{\mu \rho \eta}{k_0} V \exp\left(\frac{E_0}{RT}\right).$$

В эксперименте, проводимом при температуре  $T_e$ , аналогичный эффект будет достигнут за время

$$t_e = \frac{\mu \rho \eta}{k_0} V_e \exp\left(\frac{E_0}{RT_e}\right).$$

В качестве критерия *кинетического подобия* можно, таким образом, принять критерий

$$\pi_5 = \frac{k_0 t}{\mu \rho V \eta} \exp\left(-\frac{E_0}{RT}\right)$$

или соответствующий логарифмический критерий

$$\pi'_5 = \ln \frac{k_0 t}{\mu \rho V \eta} - \frac{E_0}{RT}.$$

Приравнивая эти критерии для натурной и экспериментальной ситуаций и принимая в качестве  $t$  гарантийный срок хранения, получаем оценку времени поверочного эксперимента

$$t_e = t \frac{V_e}{V} \exp\left(\frac{E_0}{R} \left(\frac{1}{T_e} - \frac{1}{T}\right)\right).$$

## Выводы

Проведенное исследование показало, что поверочные эксперименты по анализу химической стабильности жидкостей могут базироваться на следующих критериях подобия:

- геометрического подобия натурного и экспериментального сосудов;
- термодинамического подобия температур и гидростатических давлений в натуральных и экспериментальных условиях;
- кинетического подобия протекания химической трансформации в натуральных и экспериментальных условиях.

Указанные критерии позволяют рассчитать условия достижения заданного температурного приращения в поверочном эксперименте, обеспечивающие подобие по термодинамическому критерию, а также оценить длительность поверочного эксперимента на основе кинетического критерия.

## Список литературы

1. Седов Л. И. Методы теории подобия и размерности в механике. М.: Наука, 1965. 388 с.
2. Михеев М. А. Основы теплопередачи. М.-Л., ГЭИ, 1949. 397 с.
3. Гухман А. А. Введение в теорию подобия. М.: Высшая школа, 1973. 296 с.
4. Дьяконов Г. К. Вопросы теории подобия в области физико-химических процессов. М.: Изд-во АН СССР, 1956.
5. Darby R. Chemical engineering fluid mechanics. NY: Dekker, 2001. 559 p.
6. Моисеев А. А. Применение теории подобия в имитационном моделировании динамических процессов // Приборы и системы. 2004. № 10. С. 1.
7. Моисеев А. А. Алгебраическая интерпретация  $\pi$ -теоремы // Научные технологии в космических исследованиях Земли. 2014. № 3. С. 16.
8. Sonin A. The physical basis of dimensional analysis. Cambridge, MIT, 2001, 57 p.
9. Веников В. А. Теория подобия и моделирования. М.: Высшая школа, 1976, 479 с.
10. Сивухин Д. В. Общий курс физики. Т. 2. М.: Наука, 1990. 592 с.
11. Физические величины: справочник / Под ред. И. С. Григорьева, Е. З. Мейлихова. М.: Энергоатомиздат, 1991. 1232 с.
12. Эммануэль Н. М., Кнорре Д. Г. Курс химической кинетики. М.: Высшая школа, 1962. 415 с.

## Simulation Chemical Stability of the Liquid by the Method of Similarity Theory

*Similarity criteria widely distributed in different areas — mechanics, thermodynamics, physical chemistry etc. Here presented the building of such criteria, which are essential for liquids chemical stability in hydrostatical conditions. The following parameters considered as determinative ones: liquid's density, its molar mass and heat capacity, activation energy of transformation, vessel volume and surface area, hydrostatic pressure and liquid temperature. Since ampere and candela are ignored in frame of this consideration, only five base measurement units are used and hence only five parameters can be considered as independent. All the rest of parameters are dependent. Independent parameters choice performed using algebraic approach based on linear independency check for logarithmical representations of these parameters. Dependent criteria were built as fundamental solution elements of uniform linear system for these parameters logarithmical representations. To these criteria related the following ones: criterion of geometrical similarity of real and experimental vessels, criterion of thermodynamic similarity for temperatures and hydrostatical pressures at natural and experimental conditions, criterion of kinetic similarity of chemical transformation at natural and experimental conditions. These criteria allow determining the experiment conditions compatible with thermodynamic criterion and estimating the experiment length on kinetic criterion, which is based on Arrhenius equation.*

**Keywords:** similarity criterion, chemical stability, chemical transformation, geometrical similarity, thermodynamic similarity, kinetic similarity, temperature, hydrostatic pressure, confirmatory experiment, Arrhenius equation

### References

1. **Sedov L. I.** *Metody podobiya i razmernosti v mekhanike* [Similarity and dimensionality methods in mechanics]. Moscow, Nauka, 1965, 388 p.
2. **Micheev M. A.** *Osnovy teploperedachi* (Heat transfer fundamentals), Moscow—Leningrad, Gosenergoizdat, 1949, 397 p.
3. **Guchman A. A.** *Vvedenie v teoriyu podobiya* [Introduction to similarity theory], Moscow, Vysshaya shkola, 1973, 296 p.
4. **D'yakonov G. K.** *Voprosy teorii podobiya v oblasti fiziko-khimicheskikh processov* [Problems of similarity theory in area of physico-chemical processes], M., USSR Academy of Sciences, 1956.
5. **Darby R.** *Chemical engineering fluid mechanics*, NY, Dekker, 2001, 559 p.
6. **Moiseev A.** *Primenenie teorii podobiya v imitazionnom modelirovanii dinamicheskikh processov* [Similarity theory application in simulation of dynamical processes], *Pribory i sistemy* [Devices and systems], 2004, N 10, p. 1.
7. **Moiseev A.**  $\pi$ -theorem algebraic interpretation, H & ES Research, 2014, № 3, P. 16.
8. **Sonin A.** *The physical basis of dimensional analysis*, Cambridge, MIT, 2001, 57 p.
9. **Venikov V. A.** *Teoriya podobiya i modelirovaniya* [Similarity and modeling theory], Moscow, Vysshaya shkola, 1976, 479 p.
10. **Sivukhin D. V.** *Obshchiy kurs fiziki* [General Course of Physics]. Vol. 2. Moscow, Nauka, 1990. 592 p.
11. **Fizicheskie velichiny: spravochnik** [Physical quantities: a guide]. Ed. by Grigorev I., S. and Meylikhov E. Z. Moscow, Energoatomizdat, 1991, 1232 p.
12. **Emmanuel N. M., Knorre D. G.** *Kurs khimicheskoy kinetiki* [Course of chemical kinetics], Moscow, Vysshaya shkola, 1962, 415 p.

## Информация



22—24 октября в Центре Digital October в Москве состоится 11-я международная научно-практическая конференция "Разработка ПО/CEE-SECR 2015".

CEE-SECR — многопрофильная конференция, позволяющая специалистам и руководителям охватить широкий спектр происходящего в индустрии разработки ПО, уловить тенденции и углубить понимание, в том числе в смежных с основной специализацией областях.

В Программу конференции войдут приглашенные и конкурсные доклады, панельные дискуссии, а также открытые обсуждения.

Подробности — на сайте конференции <http://2015.secr.ru/>

УДК 004.738

**Р. Э. Асратян**, канд. техн. наук, вед. науч. сотр., rea@ipu.ru,  
**В. Н. Лебедев**, канд. техн. наук, зав. лаб., lebvini@ipu.ru,  
**В. Л. Орлов**, канд. техн. наук, вед. науч. сотр., ovl@ipu.ru  
Институт проблем управления им. В. А. Трапезникова РАН, г. Москва

## Организация защищенных каналов взаимодействия на основе применения протокола HTTPS в прокси-серверах

*Рассматривается метод организации безопасного информационного взаимодействия в распределенных системах, функционирующих в мультисетевой среде, основанный на применении протокола HTTPS и технологии прокси-серверов. Суть подхода заключается в организации защищенных HTTPS-соединений между прокси-серверами для безопасной передачи информационных запросов из одной частной локальной сети в другую через глобальную сеть с освобождением клиентских и серверных компонентов системы от функций информационной защиты.*

**Ключевые слова:** распределенные системы, Интернет-технологии, сетевые протоколы, прокси-серверы, информационное взаимодействие, информационная безопасность

### Введение

Сетевой протокол HTTPS [1] представляет собой одно из наиболее распространенных и удобных средств организации защищенного информационного взаимодействия в глобальных сетях. В последние годы значение этого протокола заметно возросло в связи с появлением возможности защиты информационных запросов к электронным сервисам в рамках технологии .NET [2, 3] на его основе (протокол SOAP "поверх" HTTPS). Этому же способствовало и широкое распространение распределенных информационных систем [4], ориентированных на работу в сложных мультисетевых средах (т. е. средах, включающих несколько связанных между собой сетей различных размеров и административной подчиненности), а также постоянно возрастающие требования к безопасности данных [5, 6].

Возможность применения HTTPS в разработках распределенных систем в значительной степени обусловлена его совместимостью с современными средствами интеграции сетей и в том числе с прокси-серверами (основы технологии HTTPS-взаимодействия через серверы-посредники предложены еще в конце 1990-х годов). Однако именно здесь разработчики распределенных систем сталкиваются с рядом проблем.

Рассмотрим мультисетевую среду, включающую несколько частных локальных сетей предприятий, объединенных глобальной сетью (рис. 1). Так как прямая IP-маршрутизация между частными сетями и глобальной сетью невозможна, задача обмена

данными между локальными сетями в данном примере решается с помощью прокси-серверов, каждый из которых оснащен двумя сетевыми интерфейсами: один из них подключен к частной сети (будем называть его "внутренним"), а другой — к глобальной сети (будем называть его "внешним"). Данный подход к решению задачи межсетового обмена часто оказывается наиболее предпочтительным по сравнению с остальными (например, с технологиями NAT [7] или VPN [7, 8]), так как на прокси-серверы можно возложить дополнительные функции по учету (регистрации), контролю, фильтрации и маршрутизации информационных запросов. В частности, контроль может включать проверку типа исходящего из частной сети HTTP-запроса (например,

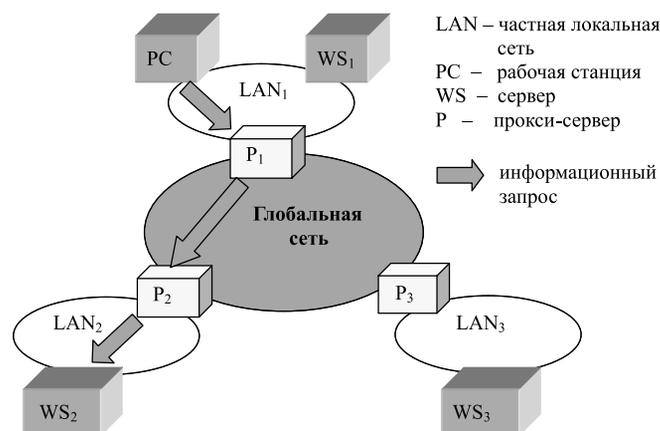


Рис. 1. Маршрутизация запросов в мультисетевой среде

допускаются только SOAP-запросы к электронным сервисам) или же проверку права доступа у источника входящего извне запроса к информационному ресурсу в частной сети.

Рассмотрим ситуацию, в которой клиентская программа, работающая в одной из частных сетей, посылает защищенный запрос к серверу, размещенному в другой частной сети по протоколу HTTPS. Каноническая форма поддержки HTTPS в прокси-сервере заключается в переключении последнего в так называемый "прозрачный режим" с помощью HTTP-метода (команды) CONNECT сразу после установления соединения с удаленным сервером. Этот режим заключается в "слепом" копировании всех данных, поступивших на внутренний сетевой интерфейс, во внешний интерфейс и наоборот вплоть до закрытия одного из соединений. Другими словами, создание защищенного канала в данном случае берут на себя клиентская и серверная программы, а от посредника требуется только "прозрачность". Так как данные от клиента и от сервера поступают в прокси-сервер в зашифрованной форме, он, очевидно, лишается возможности регистрации, контроля и вообще какой-либо обработки информационных запросов. Другими словами, в рассмотренной ситуации технология прокси-серверов теряет свое главное преимущество, например, перед технологией NAT.

В данной работе рассматривается альтернативный подход к обработке HTTPS-запросов в мульти-сетевой среде. Суть подхода заключается в организации защищенных HTTPS-тоннелей между прокси-серверами (т. е. между удаленными частными сетями) через глобальную сеть. При данном подходе удаленные клиентские и серверные программы пользуются протоколом HTTP и нисколько не заботятся об информационной защите, а на прокси-серверы возложены дополнительные функции шлюзов "HTTP → HTTPS" и "HTTPS → HTTP". Преимущество подхода заключается в том, что в данном случае прокси-серверы "видят" проходящие через них информационные запросы клиентов и ответы серверов и способны обеспечить необходимый контроль и обработку потока запросов.

### Принципы организации HTTPS-тоннеля

При организации информационного взаимодействия между удаленными частными сетями предприятий в мультисетевой среде (см. рис. 1) на первый план выходят две взаимосвязанные задачи:

- обеспечение информационной безопасности данных в условиях угроз, исходящих от глобальной сети;
- маршрутизация запросов между клиентскими и серверными программами, функционирующими в разных частных сетях.

Суть рассматриваемого подхода заключается в соединении возможностей технологии HTTPS и

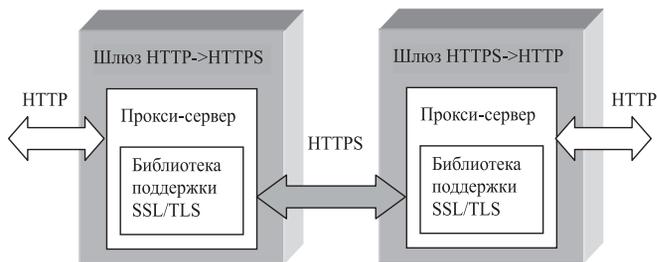


Рис. 2. Принципы организации HTTPS-тоннеля

технологии прокси-серверов для решения этих задач. Главная особенность подхода заключается в обеспечении "высокоуровневых" средств защиты и маршрутизации, основанных на заголовках и данных HTTP-запросов, а не отдельных IP-пакетов (отметим, что решение этих задач на сетевом уровне, например, на основе VPN может оказаться невозможным или совершенно неэффективным, если число частных сетей исчисляется десятками) [9].

Организация защищенного HTTPS-тоннеля с помощью двух прокси-серверов, выполняющих функции шлюзов "HTTP → HTTPS" и "HTTPS → HTTP", проиллюстрирована на рис. 2. Как видно из рисунка, каждый из прокси-серверов оснащается программной поддержкой SSL/TLS-технологии, обеспечивающей построение защищенного двустороннего канала взаимодействия "поверх" обычного TCP-соединения (отметим, что для канонического "прозрачного" режима такая поддержка не нужна). Главные особенности, которые режим шлюза привносит в поведение обычного прокси-сервера, заключаются в следующем.

- После получения HTTP-запроса от клиента, шлюз "HTTP → HTTPS" извлекает Интернет-имя (URL) адресуемого информационного ресурса из заголовка запроса и по имеющейся в его составе таблице маршрутизации определяет IP-адрес и номер порта удаленного шлюза "HTTPS → HTTP", защищающего этот ресурс (механизм маршрутизации рассмотрим чуть ниже). Кроме того, он проверяет корректность типа запроса (строка Content-type заголовка), если это необходимо.
- После установления TCP-соединения с удаленным шлюзом "HTTPS → HTTP" проводится построение двунаправленного защищенного HTTPS-канала "поверх" этого соединения с помощью программной поддержки SSL/TLS. Главное содержание этой операции заключается в согласовании методов шифрования и электронной подписи и в обмене сертификатами, содержащими открытые ключи контрагентов (в данном случае — двух шлюзов или двух организаций). С этого момента все данные, заносимые в защищенный канал, будут автоматически подписываться и шифроваться, а все данные, полученные из защищенного канала, — расшифровываться и аутентифицироваться. Сразу же после

успешного построения защищенного канала полученный от клиента информационный запрос немедленно отправляется удаленному шлюзу.

- После получения сертификата контрагента удаленный шлюз "HTTPS → HTTP" проверяет его корректность и извлекает из него имя владельца (название организации) для проверки его прав доступа к ресурсам в локальной сети. Если проверка заканчивается успешно, шлюз "HTTPS → HTTP" считывает информационный запрос по защищенному каналу, извлекает URL адресуемого информационного ресурса из заголовка запроса и по имеющейся в его составе таблице маршрутизации определяет IP-адрес и номер порта сервера в локальной сети, содержащего адресуемый ресурс. Далее следует установление TCP-соединения с этим сервером, передача запроса по новому соединению и чтение ответа сервера.
- Возврат ответа клиентской программе проводится аналогично через защищенный канал и через первичное соединение с клиентом, которые во все время обработки запроса остаются открытыми.

Как видно из изложенного, основные отличия рассмотренного подхода от канонического заключаются в следующем:

- Ни один из участников взаимодействия не использует HTTP-команду CONNECT.
- Клиентская и серверная программы не участвуют в создании защищенного канала через глобальную сеть.
- Прокси-серверы имеют возможность анализа, фильтрации и маршрутизации информационных запросов.
- Авторизацию информационных запросов и проверку прав доступа к информационным ресурсам выполняют не серверные программы, а прокси-серверы. Благодаря этому, не заслуживающие доверия запросы вообще не проникают внутрь частной сети, что повышает уровень информационной безопасности.

### Маршрутизация запросов в мультисетевой среде

Как уже отмечалось, каждый запрос, направленный из одной частной сети в другую, проходит, по крайней мере, через два прокси-сервера и через два этапа маршрутизации (см. рис. 1). Из ближайшего прокси-сервера запрос сначала передается в удаленный прокси-сервер по глобальной сети, а уже потом к адресуемому информационному ресурсу. На каждом этапе маршрутизация проводится на основе символического имени адресуемого ресурса, содержащегося в HTTP-заголовке запроса, с помощью двух таблиц маршрутизации: "глобальной", связывающей Интернет-имя ресурса с адресом удаленного прокси-сервера, и "локальной", связывающей Интернет-имя ресурса с его адресом в частной сети.

Предположим, например, что в LAN<sub>2</sub> имеется Web-сервис <http://192.168.0.20/orgdb/srv.asmx>, а в LAN<sub>3</sub> имеется Web-сервис <http://192.168.0.20/orgdb/prs.asmx> (т. е. обоим сервисам соответствует один и тот же IP-адрес). Предположим, что эти сервисы анонсированы для "внешних" пользователей под символическими именами <http://myorg.kaluga.prs.ws> и <http://myorg.omsk.prs.ws> соответственно. Для организации доступа к сервисам по этим именам в локальную таблицу маршрутизации для LAN<sub>2</sub> следует внести строку, связывающую символическое имя с URL:

```
http://myorg.kaluga.prs.ws
http://192.168.0.20/orgdb/srv.asmx,
```

а в таблицу маршрутизации для LAN<sub>3</sub> — строку:

```
http://myorg.omsk.prs.ws
http://192.168.0.20/orgdb/prs.asmx
```

Предположим, что прокси-серверу P<sub>2</sub> соответствует IP-адрес 193.232.208.2 в глобальной сети, а прокси-серверу P<sub>3</sub> — IP-адрес 193.232.208.3. В этом случае в глобальную таблицу маршрутизации должны быть добавлены две строки, связывающие символические имена с адресами ПС:

```
http://myorg.kaluga.* 193.232.208.2
http://myorg.omsk.* 193.232.208.3
```

Легко заметить, что в обеих строках использованы усеченные символические имена, оканчивающиеся символом "\*". Такая нотация обеспечивает маршрутизацию множества имен одной строкой. В данном случае, например, любой запрос к ресурсу с именем, начинающимся с <http://myorg.kaluga.>, будет перенаправлен в LAN<sub>2</sub>.

Отметим два важных свойства описанного подхода.

- Все прокси-серверы используют одну и ту же глобальную таблицу маршрутизации для обработки "исходящих" запросов.
- При изменении URL информационного ресурса в частной сети достаточно внести соответствующую коррекцию в локальную таблицу маршрутизации, чтобы никто из "внешних" клиентов не заметил этого изменения.

Совпадение IP-адресов у серверов в разных частных сетях (как в рассмотренном примере) в данном случае не создает проблем для адресации.

### Принципы функционирования шлюзов "HTTP → HTTPS" и "HTTPS → HTTP"

Наиболее специфичная черта всякого прокси-сервера заключается в том, что он объединяет в себе функциональность клиента и сервера одновременно: в отношении обслуживаемого клиента он играет роль сервера, а в отношении обслуживающего сервера — роль клиента. Это в полной мере относится и к шлюзам "HTTP → HTTPS" и "HTTPS → HTTP".

Базовая структура шлюзов проиллюстрирована на рис. 3 и рис. 4. Их функционирование подчиняется следующим общим правилам.

- Процедура Listener обеспечивает "серверное" поведение шлюза: прослушивание входящих запросов на соединение на специальном выделенном порте (системный вызов *listen*), открытие нового (первичного) TCP-соединения с клиентом (вызов *accept*) и порождение параллельной программной нити для обработки каждого входящего соединения [10]. Вся дальнейшая обработка соединения выполняется в рамках программной нити.
- Блок "Обработчик первичных соединений" обеспечивает считывание HTTP-запроса (заголовка и тела запроса) из первичного соединения, разбор заголовка запроса и извлечение URL адресуемого ресурса из тела запроса.
- Блок "Обработчик вторичных соединений" обеспечивает "клиентское" поведение шлюза: запрос и установление нового (вторичного) соединения с адресуемым сервером с помощью системных вызовов *socket* и *connect* (процедура Connector); передача HTTP-запроса серверу; прием HTTP-ответа от сервера через вторичное соеди-

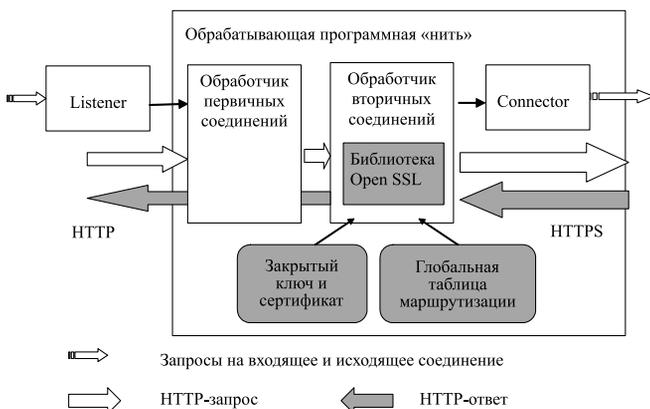


Рис. 3. Базовая структура шлюза "HTTP- > HTTPS"

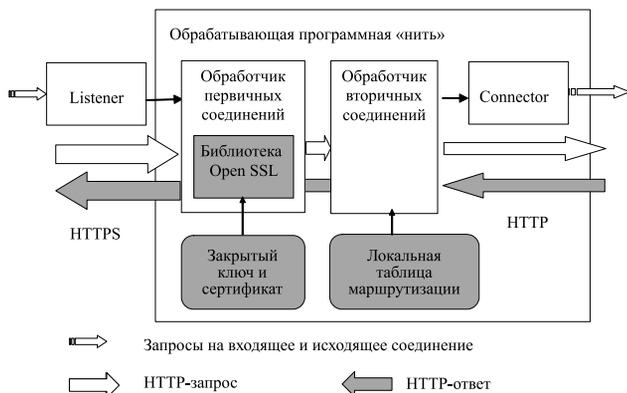


Рис. 4. Базовая структура шлюза "HTTPS- > HTTP"

нение и, наконец, передача его клиенту через первичное.

Компоненты, обеспечивающие специфическое поведение шлюзов, направленные на создание защищенного HTTPS-тоннеля, выделены на рисунках темным цветом. К их числу относятся следующие:

- библиотека функций поддержки SSL/TLS (Open SSL);
- закрытый ключ шлюза, используемый для формирования электронной подписи в исходящих данных и расшифровки входящих;
- сертификат, содержащий открытый ключ шлюза, который передается шлюзу-контрагенту при построении защищенного HTTPS-канала;
- таблица маршрутизации информационных запросов.

Как видно из изложенного, принципы функционирования шлюза "HTTP → HTTPS" и шлюза "HTTPS → HTTP" в целом схожи. Основные различия заключаются в следующем.

- В шлюзе "HTTP → HTTPS" защищенный канал надстраивается над вторичным соединением с использованием сокета, полученного с помощью системного вызова *socket*. В шлюзе "HTTPS → HTTP" защищенный канал надстраивается над первичным соединением с использованием сокета, полученного с помощью системного вызова *accept* процедурой Listener.
- Шлюз "HTTPS → HTTP" выполняет не только маршрутизацию информационных запросов к серверным программам в частной сети, но также их авторизацию и проверку прав доступа на основе данных, содержащихся в полученных сертификатах контрагентов. Эта проверка может быть реализована, например, путем добавления в каждую строку локальной таблицы маршрутизации списка владельцев сертификатов (названий организаций), которым разрешен доступ к соответствующему информационному ресурсу.

## Заключение

В таблице приведены преимущества и недостатки предлагаемого подхода в сравнении с каноническим методом поддержки HTTPS в прокси-серверах. Как видно из таблицы, основной недостаток подхода заключается в том, что в пределах частной локальной сети данные передаются в открытом виде. Поэтому его применение ограничено ситуациями, в которых частные сети можно рассматривать как "зоны доверия", а единственным источником угроз является глобальная сеть. В этих ситуациях "централизованная" защита данных в прокси-серверах, не зависящая от поведения отдельных клиентских и серверных программ, представляется оправданным решением.

Описанный подход реализован в форме многофункционального прокси-сервера, поддерживающего как канонический "прозрачный" режим, так и

### Методы поддержки протокола HTTPS в прокси-серверах

Метод	Основные преимущества	Основные недостатки	Особенности реализации
Канонический (прозрачный режим)	Защита данных на всем пути от клиента до сервера	Невозможность контроля и обработки информационных запросов в прокси-серверах	Не требуется использования какой-либо программной поддержки SSL/TLS в прокси-серверах
HTTPS-тоннель	Возможность контроля и обработки информационных запросов в прокси-серверах	Защита данных только при информационном обмене между прокси-серверами (т.е. только в глобальной сети). Внутри частной сети данные не защищены	Требуется использование библиотек поддержки SSL/TLS в прокси-серверах

режимы шлюзов "HTTP → HTTPS" и "HTTPS → HTTP" (реализация проведена на языке C++ с использованием библиотеки OpenSSL в качестве программной поддержки технологии SSL/TLS). Опыт работы с ним показал, что описанный подход является достаточно общим: однажды созданная система защищенных тоннелей межсетевое взаимодействие может быть использована несколькими совершенно различными распределенными системами, работающими в мультисетевой среде.

#### Список литературы

1. Кришнамурти Б., Рэксфорд Д. Web протоколы. М.: Бинум, 2010. 592 с.
2. Шапошников И. В. Web-сервисы Microsoft .NET. СПб.: БХВ-Петербург, 2002. 336 с.
3. Мак-Дональд М., Шпуста М. Microsoft ASP.NET 3.5 с примерами на C# 2008 и Silverlight 2 для профессионалов. М.: Вильямс, 2009. 1408 с.
4. Танненбаум Э., Ван Стен М. Распределенные системы. Принципы и парадигмы. СПб.: Питер, 2003, 877 с.
5. Згоба А. И., Маркелов Д. В., Смирнов П. И. Кибербезопасность: угрозы, вызовы, решения // Вопросы кибербезопасности. 2014. № 5. С. 30—38.
6. Щеглов А. В. Защита компьютерной информации от несанкционированного доступа. СПб.: Наука и техника, 2004. 384 с.
7. Хант К. TCP/IP. Сетевое администрирование. СПб.: Питер, 2007. 816 с.
8. Андреев А. Г., Беззубов Е. Ю., Емельянов М. М., Кокорева О. И., Чекмарев А. Н. Windows 2000: Server и Professional. СПб.: БХВ-Санкт-Петербург, 2000. 1056 с.
9. Асратян Р. Э., Лебедев В. Н. Применение технологии прокси-серверов в распределенных системах // Информационные технологии. 2013. № 6. С. 7—11.
10. Снейдер Й. Эффективное программирование TCP/IP. Библиотека программиста. СПб.: Символ-Плюс, 2002. 320 с.

R. E. Asratian, Leading Researcher, rea@ipu.ru, V. N. Lebedev, Head of Laboratory, lebvini@ipu.ru, V. L. Orlov, Leading Researcher, ovl@ipu.ru, Institute of Control Science RAS

## Creating a Secure Communication Channels Based on the use of the HTTPS in Proxy-Servers

*The method of organization of secure interactions in distributed systems, operating in multi-network environment, is considered. The approach is based on the use of HTTPS protocol and proxy-server technology. The essence of approach consists in the organization of the protected HTTPS connections between proxy servers for safe transfer of information queries from one private local network in another through a global network while client and server components of distributed system are released from of information protection functions. In this approach client and server programs use HTTP protocol and proxy servers have additional functions of "HTTP- > HTTPS" and "HTTPS- > HTTP" gateways. The main advantage of approach over known (for example, VPN) consists in providing "high-level" means of HTTP queries routing and filtration the based on headers of HTTP queries, but not on separate IP packages. The application of approach is limited to situations in which private local networks can be considered as "a trust zone", and the only source of threats is the global network.*

**Keywords:** distributed systems, Internet technologies, network protocols, proxy servers, remote interactions, data security

#### References

1. Krishnamurti B., Reksford D. Web protokoly. Moscow, Binom, 2010, 592 p.
2. Shaposhnikov I. V. Web-servisy Microsoft .NET. SPb.: BHV-Peterburg, 2002. 336 p.
3. Mak-Donal'd M., Shpushhta M. Microsoft ASP.NET 3.5 s primerami na C# 2008 i Silverlight 2 dlja professionalov. Moscow, Vil'jams, 2009, 1408 p.
4. Tannenbaum E., Van Sten M. Raspredelelennye sistemy. Principy i paradigmy. SPb.: Piter, 2003, 877 p.
5. Zgoba A. I., Markelov D. V., Smirnov P. I. Kiberbezopasnost': ugrozy, vyzovy, reshenija, Voprosy kiberbezopasnosti, 2014, no. 5, pp. 30—38.
6. Shheglov A. V. Zashhita komp'yuternoj informacii ot nesankcionirovannogo dostupa Sankt-Peterburg, Nauka i tehnika, 2004, 384 p.
7. Hant K. TCP/IP. Setevoe administrirovanie, Sankt-Peterburg, Piter, 2007. 816 p.
8. Andreev A. G., Bezzubov E. Ju., Emel'janov M. M., Kokoreva O. I., Chekmarev A. N. Windows 2000: Server i Professional. SPb.: BHV-Sankt-Peterburg, 2000, 1056 p.
9. Asratian R. E., Lebedev V. N. Primenenie tehnologii proksi-serverov v raspredelelennyh sistemah, Informacionnye tehnologii, 2013, no. 6, pp. 7—11.
10. Snejder J. Effektivnoe programmirovanie TCP/IP. Biblioteka programmista, Sankt-Peterburg, Simvol-Pljus, 2002, 320 p.

А. Э. Саак, д-р техн. наук, доц., e-mail: saak@tgn.sfedu.ru,  
Южный федеральный университет, г. Таганрог

## Диспетчеризация массивов заявок суммарной ресурсной меры, равной квадрату целого числа

В дальнейшем развитии среды ресурсных прямоугольников определяются массивы заявок с суммарной ресурсной мерой, равной квадрату целого числа, называемые имеющими точную ресурсную меру. Для указанных массивов заявок пользователей проводится диспетчеризация и вычисляются эвристические меры ресурсных оболочек начально-кольцевого, уровневого, углового уровневого и алгоритма последовательных приближений. Для этих полиномиальных алгоритмов показывается сохранение свойства адаптированности на массивах заявок, имеющих точную ресурсную меру.

**Ключевые слова:** линейные полиэдры точной ресурсной меры, линейные полиэдры кругового типа, Grid-система, диспетчеризация, неэвклидова эвристическая мера, полиномиальная трудоемкость алгоритма, начально-кольцевой алгоритм, уровневый алгоритм, угловой уровневый алгоритм, алгоритм последовательных приближений

### Введение

Для массивов заявок пользователей Grid-систем в предыдущих работах автора [1–5] определена среда ресурсных прямоугольников, предложена и разработана классификация множества заявок на круговой, гиперболический и параболический квадратичные типы. Там же предложены и исследованы полиномиально трудоемкие алгоритмы распределения ресурсов, адаптированные под круговой тип массива заявок. Оптимальная укладка натуральной последовательности квадратов по критерию минимума площади объемлющего квадрата содержит пустоты [6, 7]. В статье выделяется класс массивов заявок с суммарной ресурсной мерой, равной квадрату целого числа, и имеющих безпустотную ресурсную оболочку при оптимальной укладке. Исследуется адаптированность начально-кольцевого, уровневого, углового уровневого и алгоритма последовательных приближений для выделенного класса.

### Массив заявок, имеющий точную ресурсную меру

При представлении заявки пользователя для обслуживания диспетчером Grid-системы ресурсным прямоугольником горизонтальное и вертикальное измерения, соответственно, принимаются равными числу единиц ресурса времени и процессоров, требуемому для обработки. Символом  $a(j) \times b(j)$  обозначается  $j$ -я заявка, требующая  $a(j)$  единиц времени и  $b(j)$  единиц процессоров. Множество заявок пользователей представляется последовательностью ресурсных прямоугольников — линейной полиэдралью (рис. 1) [8].

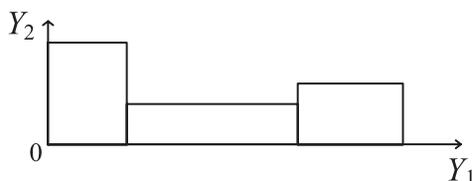


Рис. 1. Линейная полиэдраль заявок пользователей

Линейная полиэдраль ресурсных прямоугольников, сумма площадей которых является квадратом целого числа, называется полиэдралью, имеющей точную ресурсную меру. Примером могут служить линейные полиэдры элементов парtridge (от англ. *partridge*), содержащие один квадрат  $1 \times 1$  со стороной 1, два квадрата  $2 \times 2$  со стороной 2, ...,  $k$  квадратов  $k \times k$  со стороной  $k$ , согласно равенству [9]

$$\sum_{i=1}^k i \cdot i^2 = \left( \sum_{i=1}^k i \right)^2 = \left( \frac{k(k+1)}{2} \right)^2.$$

В случае когда такая линейная полиэдраль укладывается в соответствующий квадрат со стороной  $k(k+1)/2$ , это называется парtridgeной облицовкой (от англ. *partridge tiling*) [10], а соответствующее  $k$  — числом парtridge квадрата (от англ. *partridge number of the square*) [11]. Как было показано в работах [12–14], парtridgeными облицовками являются линейные полиэдры элементов парtridge при  $k = 11, 12, 13, 14$ . Отметим, что все четыре линейные полиэдры, индуцированные парtridgeными облицовками, относятся к круговому типу.

### Диспетчеризация массива заявок, имеющего точную ресурсную меру

Для Grid-систем с централизованной структурой системы диспетчеризации и коаллокацией ресурсов (мультисайтное диспетчеризование) в качестве графической интерпретации была предложена мо-

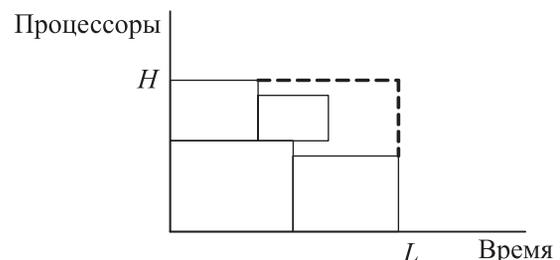


Рис. 2. Ресурсная оболочка заявок пользователей

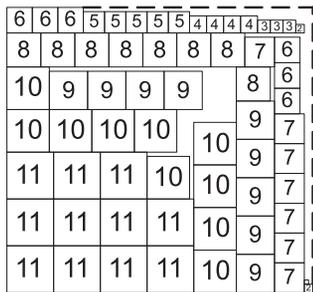


Рис. 3. Укладка линейной полиэдры I начально-кольцевым алгоритмом

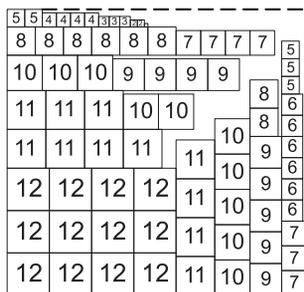


Рис. 4. Укладка линейной полиэдры II начально-кольцевым алгоритмом

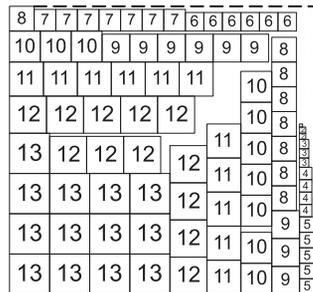


Рис. 5. Укладка линейной полиэдры III начально-кольцевым алгоритмом

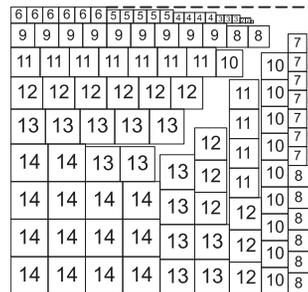


Рис. 6. Укладка линейной полиэдры IV начально-кольцевым алгоритмом

Таблица 1

Эвристические меры ресурсных оболочек начально-кольцевого алгоритма

Номер линейной полиэдры	Эвристическая мера	$\Delta$ , %
I	0,56	12
II	0,57	14
III	0,56	12
IV	0,54	8

дель ресурсного квадранта [5, 8]. Качество диспетчирования эвристических алгоритмов оценивается неевклидовой эвристической мерой, учитывающей наряду с площадью и форму занятую ресурсной области

$$\frac{1}{2} \left( \frac{LH + (L - H)^2}{k - 1} + \sum_{j=0}^k a(j)b(j) \right),$$

где  $L$  — протяженность и  $H$  — уровень по вертикали ресурсной оболочки (рис. 2) [5].

Эвристическая мера достигает минимума, равного  $1/2$ , при безпустотной укладке в квадрат. Указанные четыре линейные полиэдры элементов партриджа являются такими укладками. Вычислим эвристические меры ресурсных оболочек, получаемых при диспетчировании этих линейных полиэдров полиномиальными алгоритмами [5]: началь-

но-кольцевым, уровневым, угловым уровневым и алгоритмом последовательных приближений.

Линейные полиэдры элементов партриджа обозначим следующим образом: для  $k = 11$  — линейная полиэдраль I, для  $k = 12$  — линейная полиэдраль II, для  $k = 13$  — линейная полиэдраль III, для  $k = 14$  — линейная полиэдраль IV.

Построения начально-кольцевым алгоритмом для линейных полиэдров I—IV приведены на рис. 3—6.

Эвристические меры ресурсных оболочек начально-кольцевого алгоритма и погрешность  $\Delta$  в % относительно оптимального значения, равного  $1/2$ , приведены в табл. 1.

Заметим, что эвристические меры ресурсных оболочек начально-кольцевого алгоритма не превосходят значения  $1/2 + 0,07$ .

Построения уровневым алгоритмом для линейных полиэдров I—IV приведены на рис. 7—10.

Таблица 2

Эвристические меры ресурсных оболочек уровневого алгоритма

Номер линейной полиэдры	Эвристическая мера	$\Delta$ , %
I	0,55	10
II	0,56	12
III	0,55	10
IV	0,57	14

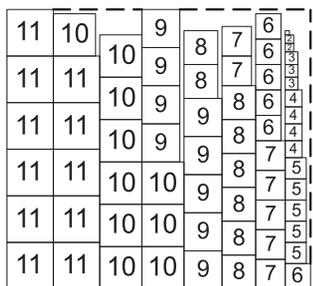


Рис. 7. Укладка линейной полиэдры I уровневым алгоритмом

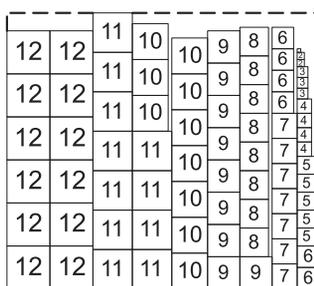


Рис. 8. Укладка линейной полиэдры II уровневым алгоритмом

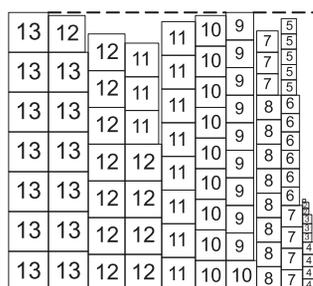


Рис. 9. Укладка линейной полиэдры III уровневым алгоритмом

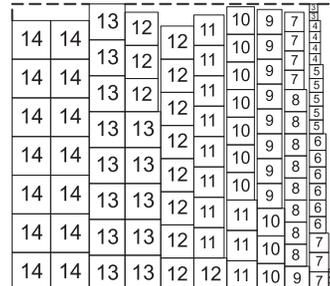


Рис. 10. Укладка линейной полиэдры IV уровневым алгоритмом

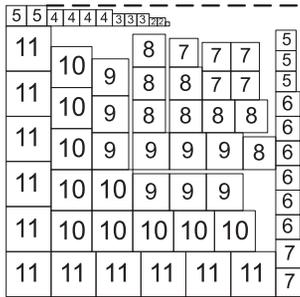


Рис. 11. Укладка линейной полиэдрала I угловым уровнем алгоритмом

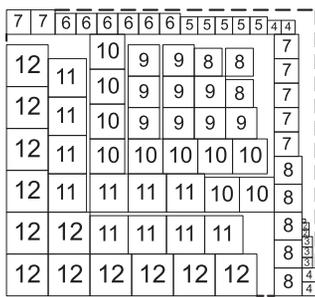


Рис. 12. Укладка линейной полиэдрала II угловым уровнем алгоритмом

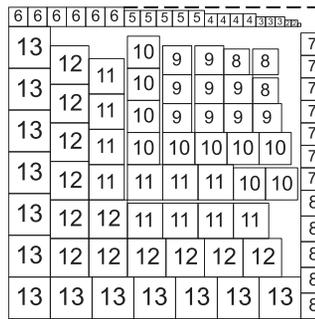


Рис. 13. Укладка линейной полиэдрала III угловым уровнем алгоритмом

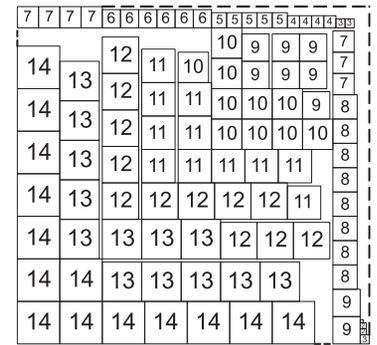


Рис. 14. Укладка линейной полиэдрала IV угловым уровнем алгоритмом

Таблица 3

Эвристические меры ресурсных оболочек углового уровня алгоритма

Номер линейной полиэдрала	Эвристическая мера	$\Delta$ , %
I	0,60	20
II	0,60	20
III	0,58	16
IV	0,59	18

Эвристические меры ресурсных оболочек углового алгоритма и погрешность  $\Delta$  в % относительно оптимального значения, равного  $1/2$ , приведены в табл. 2.

Заметим, что эвристические меры ресурсных оболочек углового алгоритма не превосходят значения  $1/2 + 0,07$ .

Построения угловым уровнем алгоритмом для линейных полиэдралей I—IV приведены на рис. 11—14.

Эвристические меры ресурсных оболочек углового алгоритма и погрешность  $\Delta$  в % относительно оптимального значения, равного  $1/2$ , приведены в табл. 3.

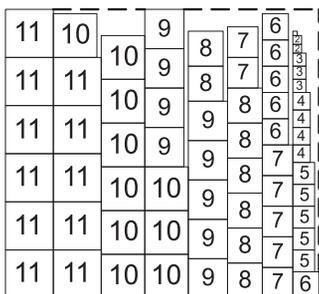


Рис. 15. Укладка линейной полиэдрала I алгоритмом последовательных приближений

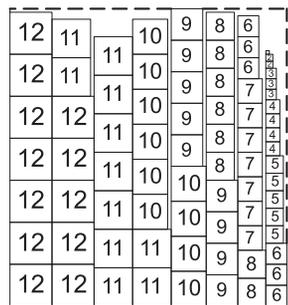


Рис. 16. Укладка линейной полиэдрала II алгоритмом последовательных приближений

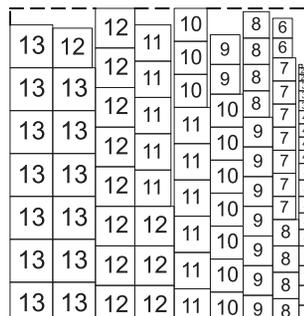


Рис. 17. Укладка линейной полиэдрала III алгоритмом последовательных приближений

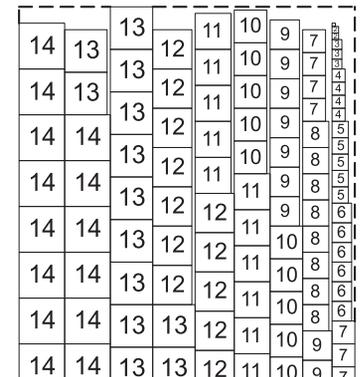


Рис. 18. Укладка линейной полиэдрала IV алгоритмом последовательных приближений

Заметим, что эвристические меры ресурсных оболочек углового уровня алгоритма не превосходят значения  $1/2 + 0,1$ .

Построения алгоритмом последовательных приближений для линейных полиэдралей I—IV приведены на рис. 15—18.

Эвристические меры ресурсных оболочек алгоритма последовательных приближений и погрешность  $\Delta$  в % относительно оптимального значения, равного  $1/2$ , приведены в табл. 4.

Заметим, что эвристические меры ресурсных оболочек алгоритма последовательных приближений не превосходят значения  $1/2 + 0,06$ .

Таблица 4

Эвристические меры ресурсных оболочек алгоритма последовательных приближений

Номер линейной полиэдрала	Эвристическая мера	$\Delta$ , %
I	0,55	10
II	0,55	10
III	0,55	10
IV	0,56	12

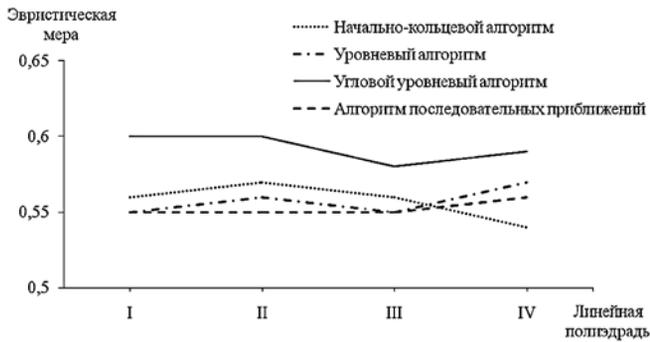


Рис. 19. Эвристические меры ресурсов оболочек начально-кольцевого, уровневого, углового уровневого и алгоритма последовательных приближений

Графики эвристической меры ресурсов оболочек начально-кольцевого, уровневого, углового уровневого и алгоритма последовательных приближений диспетчеризации линейными полиэдрами, имеющими точную ресурсную меру, кругового типа I—IV показаны на рис. 19.

Видим, что начально-кольцевой и алгоритм последовательных приближений имеют меньшую эвристическую меру ресурсов оболочек. Целесообразность использования предложенных алгоритмов при диспетчеризации линейными полиэдрами, имеющими точную ресурсную меру, кругового типа подтверждается значением отклонения, не превышающим 20 % от оптимального значения.

### Заключение

В развитии среды ресурсов прямоугольников дается определение массивов заявок, имеющих точную ресурсную меру. Приводятся примеры такие, как массивы, содержащие один квадрат со стороной 1, два квадрата со стороной 2, ...,  $k$  квадратов со стороной  $k$ . Для указанных массивов вычисляются эвристические меры ресурсов оболочек начально-кольцевого, уровневого, углового уровневого и алгоритма последовательных приближений. Показывается адаптированность исследуемых полиномиальных алгоритмов для массивов заявок кругового типа, имеющих точную ресурсную меру. Так, для укладки массива 105 квадратов, соответствующего  $k = 14$ , время работы оптимального алгоритма на одноядерном компьютере 3.3 GHz Intel

Xeon превышает 13 дней и 8 ч, тогда как начально-кольцевой алгоритм практически мгновенно дает решение с погрешностью 8 %.

### Список литературы

1. Саак А. Э. Локально-оптимальные ресурсные распределения // Информационные технологии. 2011. № 2. С. 28—34.
2. Саак А. Э. Алгоритмы диспетчеризации в Grid-системах на основе квадратичной типизации массивов заявок // Информационные технологии. 2011. № 11. С. 9—13.
3. Саак А. Э. Диспетчеризация в Grid-системах на основе однородной квадратичной типизации массивов заявок пользователей // Информационные технологии. 2012. № 4. С. 32—36.
4. Саак А. Э. Сравнительный анализ полиномиальных алгоритмов диспетчеризации в Grid-системах // Информационные технологии. 2012. № 9. С. 28—32.
5. Саак А. Э. Полиномиальные алгоритмы распределения ресурсов в Grid-системах на основе квадратичной типизации массивов заявок // Информационные технологии. 2013. № 7. Приложение. 32 с.
6. Korf R. Optimal rectangle packing: New results // Proceedings of the fourteenth international conference on automated planning and scheduling (ICAPS 2004). Whistler, British Columbia, Canada, June 3—7, 2004. P. 142—149.
7. Korf R., Moffitt M., Pollack M. Optimal rectangle packing // Annals of Operations Research. 2010. Vol. 179, N. 1. P. 261—295.
8. Саак А. Э. Управление ресурсами и заявками пользователей в Grid-системах с централизованной архитектурой // Тр. XII Всероссийского совещания по проблемам управления ВСПУ—2014. Москва, 16—19 июня 2014 г. М.: Институт проблем управления им. В. А. Трапезникова РАН, 2014. С. 7489—7498.
9. Ahearn S., Jepsen C. On tiling an  $m \times m$  square with  $m$  squares // Crux Mathematicorum. 1993. Vol. 19, N. 7. P. 189—191.
10. Friedman E. Reptiles, Partridges, and Golden Bees: Tiling Shapes with Similar Copies. 2003. URL: [www2.stetson.edu/~efriedman/papers/reptiles.ppt](http://www2.stetson.edu/~efriedman/papers/reptiles.ppt).
11. Hamlyn P., Friedman E. Partridge numbers // Geombinatorics Quarterly. 2003. Vol. XIII, Is. 1. P. 10—18.
12. Simonis H., O'Sullivan B. Search strategies for rectangle packing. 2008. URL: <http://4c.ucc.ie/~hsimonis/cp2008slides.pdf>.
13. Ågren M., Beldiceanu N., Carlsson M., Sbihi M., Truchet C., Zampelli S. Six ways of integrating symmetries within non-overlapping constraints. In W.-J. van Hoeve and J. N. Hooker, ed., CPAIOR. // LNCS. 2009. Vol. 5547. P. 11—25.
14. Hougardy S. A Scale Invariant Exact Algorithm for Dense Rectangle Packing Problems. Research Institute for Discrete Mathematics, University of Bonn, Report No: 101020, 2012.

## Scheduling of Sets of Multiprocessor Tasks of the Aggregate Resource Measure Equaling the Square of an Integer

In the paper it is given further investigation of the resource rectangles environment, which represents formal instrumentation of time and computer resources scheduling. For a model of the set of multiprocessor tasks in the form of extensive linear polyhedral of resource rectangles and for a model of the Grid system with centralized architecture which supports multisite task performance in the form of resource quadrant, author solves the problem of allocation with minimization of a resource enclosure heuristic measure. The Non-Euclidean heuristic measure which takes into consideration both area and the form of an occupied resource region, reaches its minimum value at in-square packing without emptiness. In the paper we denote the linear polyhedrals which have precise resource measure and have a sum of resource rectangles area which is one of integral square. The class of linear polyhedrals which are induced by the elements of partridge, so named partridge tiling, is considered. The analysis of the linear polyhedrals which are induced by partridge tiling showed that they belong to circular type. Scheduling was performed and the heuristic measure indicators for the resource enclosures created by an initial ring algorithm, level algorithm, angular-level algorithm and successive approximation algorithm were calculated. We prove that the polynomial algorithms mentioned retain their adaptedness characteristic when used for circular-type linear polyhedrals of precise resource measure. The comparative analysis of the algorithms shows that the initial ring algorithm and successive approximation algorithm indicate smaller heuristic measure of the resource enclosures than the others. Taking into consideration quality of the resource enclosure filling in comparison with time of optimal scheduling, we can see advantages of the algorithms, considered in the paper. For example the time of optimal square packing of 105 squares of a linear polyhedral which is induced by partridge tiling with partridge number of the square equaled 14 on PC 3.3 GHz Intel Xeon with one CPU core exceed 13 days and 8 hours while the initial ring algorithm performs square packing almost instantly with 8 % deviation of the heuristic measure indicator from optimal one.

**Keywords:** linear polyhedrals of precise resource measure, circular-typed linear polyhedrals, partridge tiling, Grid system, resource management, scheduling, Non-Euclidean heuristic measure, polynomial complexity of an algorithm, initial ring algorithm, level algorithm, angular-level algorithm, successive approximation algorithm

### References

1. Saak A. Eh. Lokalno-optimalnye resursnye raspredeleniya. *Informatsionnye tekhnologii*. 2011, no. 2, pp. 28–34.
2. Saak A. Eh. Algoritmy dispetcherizatsii v Grid-sistemakh na osnove kvadrachnoy tipizatsii massivov zayavok. *Informatsionnye tekhnologii*, 2011, no. 11, pp. 9–13.
3. Saak A. Eh. Dispetcherizatsiya v GRID-sistemakh na osnove odnorodnoy kvadrachnoy tipizatsii massivov zayavok polzovateley, *Informatsionnye tekhnologii*. 2012, no. 4, pp. 32–36.
4. Saak A. Eh. Sravnitelniy analiz polinomialnykh algoritmov dispetcherizatsii v GRID-sistemakh. *Informatsionnye tekhnologii*, 2012, no. 9, pp. 28–32.
5. Saak A. Eh. Polinomialnye algoritmy raspredeleniya resursov v Grid-sistemakh na osnove kvadrachnoy tipizatsii massivov zayavok, *Informatsionnye tekhnologii*, 2013, no. 7. Prilozhenie. 32 p.
6. Korf R. Optimal rectangle packing: New results. *Proceedings of the fourteenth international conference on automated planning and scheduling (ICAPS 2004)*. Whistler, British Columbia, Canada, June 3–7, 2004, pp. 142–149.
7. Korf R., Moffitt M., Pollack M. Optimal rectangle packing. *Annals of Operations Research*, 2010, vol. 179, no. 1, pp. 261–295.
8. Saak A. Eh. Upravleniye resursami i zayavkami polzovateley v Grid-sistemakh s centralizovannoy arkhitekturoy, *Trudy XII Vserossiyskogo soveshanya po problem upravleniya VSPU–2014*. Moskva, 16 ijunya — 19 ijunya 2014 g. Moscow: Institut problem upravleniya im. V. A. Trapeznikova RAN. 2014, pp. 7489–7498.
9. Ahearn S., Jepsen C. On tiling an  $m \times m$  square with  $m$  squares. *Crux Mathematicorum*. 1993, vol. 19, no. 7, pp. 189–191.
10. Friedman E. *Reptiles, Partridges, and Golden Bees: Tiling Shapes with Similar Copies*. URL: [www2.stetson.edu/~efriedma/papers/reptiles.ppt](http://www2.stetson.edu/~efriedma/papers/reptiles.ppt).
11. Hamlyn P., Friedman E. Partridge numbers. *Geombinatorics Quarterly*. 2003, vol. XIII, Is. 1, pp. 10–18.
12. Simonis H., O’Sullivan B. Search strategies for rectangle packing. 2008. URL: <http://4c.ucc.ie/~hsimonis/cp2008slides.pdf>.
13. Ågren M., Beldiceanu N., Carlsson M., Sbihi M., Truchet C., Zampelli S. Six ways of integrating symmetries within non-overlapping constraints. In W.-J. van Hoeve and J. N. Hooker, ed., CPAIOR. *LNCS*, 2009, vol. 5547, pp. 11–25.
14. Hougardy S. *A Scale Invariant Exact Algorithm for Dense Rectangle Packing Problems*. Research Institute for Discrete Mathematics, University of Bonn, Report No: 101020, 2012.

# ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ И ИЗОБРАЖЕНИЙ DIGITAL PROCESSING OF SIGNALS AND IMAGES

УДК 004.042

**В. П. Потапов**, д-р техн. наук, зам. директора — директор филиала, potapov@ict.sbras.ru,

**С. Е. Попов**, канд. техн. наук, ст. науч. сотр., popov@ict.sbras.ru

Федеральное государственное бюджетное учреждение науки Институт вычислительных технологий  
Сибирского отделения Российской академии наук, Новосибирск

## Многопоточный алгоритм конвертации чередования данных каналов гиперспектральных снимков

*Представлен высокопроизводительный алгоритм преобразования данных спектральных каналов (Band Interleave Conversion) сенсора EO-1 Hyperion, предусматривающий возможность запуска его на многопроцессорных платформах в мультипоточном режиме и обеспечивающий эффективное выполнение на низкопроизводительных системах ввода/вывода. Предложена реализация порционного считывания данных снимка в оперативную память с последующим размещением значений (Digitals Number) в целевых буферных массивах в несколько потоков, рассчитанных по числу спектральных каналов или по числу линий снимка. Представлено также расширение для программного продукта Exelis ENVI, реализующее разработанный алгоритм, на базе технологии GUI-WIDGETS в интеграции с пакетами Java SwingX. Приведены результаты тестирования представленных алгоритмов с их базовыми аналогами программного комплекса Exelis ENVI.*

**Ключевые слова:** преобразование порядка следования данных, спектральные каналы, гиперспектральные изображения, многопоточность, Java, IDL-Bridge, EO-1 Hyperion

### Введение

Данные зондирования Земли поступают в виде изображений, как правило, в цифровой форме, обработка ведется на ЭВМ, поэтому технологии дистанционного зондирования тесно связаны с цифровой обработкой изображений. Предварительная обработка данных дистанционного зондирования заключается в геометрической коррекции спутниковых изображений, радиометрической и атмосферной коррекции, восстановлении пропущенных пикселей и улучшении изображений путем изменения контраста. Одним из наиболее важных этапов предварительной обработки космических снимков является атмосферная коррекция. В настоящее время большая часть данных, попадающих в руки конечного пользователя, уже атмосферно скорректирована. Однако в связи с тем, что процесс коррекции больших объемов данных полностью автоматизирован, используемые для коррекции алгоритмы рассчитывают исходя из усредненных показателей и не учитывают особенности получения каждого конкретного снимка, а также возможность оптимизации вычислительного процесса для каждого конкретного программного продукта или отдельных сторонних специализированных модулей.

Существует ряд работ, направленных на оптимизацию алгоритмов пред- и постобработки мульти- и гиперспектральных изображений, а также на сравнение разрешающей способности алгоритмов. Так, в работе [10] представлен алгоритм *Multi-Scale*

*Retinex (MSR)*, обеспечивающий сжатие динамического диапазона, снижение зависимости от условий освещения, а также улучшение пространственного разрешения мультиспектральных данных. В частности, результаты моделирования показывают, что, несмотря на степень снижения качества изображения вследствие изменения атмосферного излучения, ошибки классификации могут быть существенно уменьшены путем предварительной обработки данных изображения с помощью MSR. В работе [8] сравниваются алгоритмы атмосферной коррекции *Atmospheric CORrection Now (ACORN)*, *the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH)*, and *ATmospheric CORrection (ATCOR 2-3)* для гиперспектральных данных сенсора EO-1 Hyperion для кросс-корреляции и приводятся особенности поглощения в определенных длинах волн. Показано, что в зависимости от различных диапазонов длин волн алгоритм атмосферной коррекции ACORN показал лучшую корректирующую способность для снимков EO-1 Hyperion для определения литологии и минералогического состава природных земельных геоматериалов. В работе [7] описаны результаты использования моделей алгоритмов ATREM, HATCH, ACORN и FLAASH и приведены оценки коэффициентов отражения на верхней границе атмосферы для водяного пара. Показано, что существуют значительные различия между моделями атмосферной коррекции протестированных алгоритмов. Отличия моделей равны или больше, чем ошибки, вносимые калибровкой длин волн или

шумами сенсоров. Этот результат делает невозможным узнать, какие модели работают лучше. В обоих случаях (по данным AVIRIS и Hyperion) графики коэффициентов отражения, полученные из сцен с наиболее высокими показателями для водяного пара, содержали значительное число артефактов и отклонение от сглаженной кривой отражения. Обнаружено, что использование алгоритма MATCH-2D с опцией *"column-by-column wavelength calibration"*, позволило получить более "сглаженные" спектры отражения, чем при использовании метода *"average spectral calibration"* для всего массива данных. Однако недостатком данного подхода явилось значительное увеличение времени вычислений.

Не менее важными являются вопросы, связанные с вычислительной производительностью алгоритмов пред- и постобработки спектральных изображений. В работах [1] и [6] рассматривается программный прототип системы автоматизированного, высокоточного алгоритма с малой задержкой для расчета атмосферной коррекции на основе C++ для модуля FLAASH. Предложенный в работах [1] и [6] подход заключается в замене расчетов моделей переноса радиации модуля MODTRAN на предварительно рассчитанные справочные таблицы, используя для этого параллельную обработку на многоядерных/многопроцессорных вычислительных системах.

В настоящее время на рынке программных продуктов есть готовые решения для визуализации и обработки данных дистанционного зондирования Земли (ДЗЗ), которые включают в себя набор инструментов для проведения полного цикла обработки данных от ортотрансформирования и пространственной привязки изображения до получения необходимой информации и ее интеграции с данными ГИС. Однако программная реализация алгоритмов, интегрированных в такой софт, не оптимизирована, требует достаточно много машинных ресурсов и длительного времени на обработку изображений. Например, работа модуля *Band Interleave* с опцией подготовки данных для атмосферной коррекции алгоритмами FLAASH в программном комплексе ENVI для гиперспектрального снимка (7,7 × 185 км) сенсора EO-1 составляет около 2,5 ч. При этом некорректно используются ресурсы файловой подсистемы вследствие неоптимизированного порядка считывания файлов данных спектральных каналов. Конечно, стоит сказать, что в том же программном продукте Exelis ENVI есть встроенная технология мультипоточности (The IDL Thread Pool: [сайт]. URL: [http://www.exelisvis.com/docs/The\\_Thread\\_Pool.html](http://www.exelisvis.com/docs/The_Thread_Pool.html)), но она представлена по принципу *hard coded*, т. е. разработчик ее не может настраивать и ею управлять, а самое главное, она реализована только для простейших математических операций и создания массивов, что не позволяет в полной мере применять технологию объектно-ориентированного подхода Java.

В данной работе описана разработка высокопроизводительного алгоритма для предваритель-

ной обработки (подготовки) гиперспектральных данных с последующей их передачей на вход модуля атмосферной коррекции программного продукта Exelis ENVI.

## Постановка задачи

Разработать алгоритм преобразования порядка следования данных отдельных каналов, их программную реализацию в виде расширений для программного комплекса ENVI с использованием технологии IDL-Java Bridge. Программный функционал алгоритмов должен предусматривать возможность запуска их на многопроцессорных платформах в мультипоточном режиме и обеспечивать эффективное выполнение на низкопроизводительных системах ввода/вывода, с объемом оперативной памяти до 2 Гбайт.

## Используемые данные и программные продукты

Для тестирования разрабатываемых алгоритмов использованы данные гиперспектральных изображений, полученных с ресурса *EarthExplorer* (EarthExplorer: [сайт]. URL: <http://earthexplorer.usgs.gov/>) Геологической службы США. В работе рассматривались изображения сенсора EO-1 Hyperion с полосами съемки 7,7 × 42 км и 7,7 × 185 км. Гиперспектральный сенсор Hyperion предназначен для получения цифровых изображений земной поверхности с пространственным разрешением до 30 м в 242 спектральных диапазонах видимой и инфракрасной областей.

Гиперспектральные снимки сенсора EO-1 Hyperion уровня L1T представляют собой набор из 242 файлов, соответствующих спектральным каналам в графическом формате GeoTIFF, а также файл метаданных с расширением (.TXT или .MET). Файлы имеют нативный формат именования, соответствующий схеме EOlsppprrrYYYYDDDXXML\_BBBB\_TTT.TIF (Department of the Interior U. S. Geological Survey, 2006). Размер снимка определяется числом линий (параметр PRODUCT\_LINES) и числом пикселей одной линии (параметр PRODUCT\_SAMPLES). Эта информация размещается в файле метаописания (EarthExplorer: [сайт]. URL: <http://earthexplorer.usgs.gov/>).

Каждый TIFF-файл спектрального канала содержит теги с информацией о типе, количестве и размере данных, в нем размещенных. В работе были использованы стандартные теги TIFF-формата:

- StripOffsets — содержит массив значений смещений (по одному на линии), которые указывают позицию первого байта каждой линии в файле TIFF; первый элемент массива указывает смещение первой линии, второй — смещение второй линии и т. д.
- StripByteCount — описывает массив значений, указывающих размер каждой линии в байтах.

Гиперспектральные снимки, загружаемые с ресурса EarthExplorer (EarthExplorer: [сайт]. URL: <http://earthexplorer.usgs.gov/>), имеют формат преобразования порядка следования данных каналов BSQ (*band sequential*) (рис. 1, см. четвертую сторону обложки). Данный формат хранит информацию о каждом спектральном канале изображения "за раз". Другими словами, данные для всех пикселей первого канала расположены первыми (с нулевым смещением), затем идут данные всех пикселей второго канала со смещением (число каналов)  $\times$  (число линий)  $\times$  (число пикселей в линии)  $\times 2$  и так далее по числу спектральных каналов. Тип данных для каждого значения пикселя — Short, размером 16 бит.

Высокопроизводительные расчетные алгоритмы разработаны с применением объектно-ориентированного языка Java, расширения для программного комплекса Exelis ENVI созданы на базе языка IDL, связующим интерфейсом послужила технология Java-IDL Bridge.

### Алгоритм и его программная реализация

Основной акцент в работе ставился на создание алгоритма препроцессинга (преобразование порядка следования данных отдельных каналов *Band Interleave Conversion*) полного снимка с данными уровня LIT без выделения на нем областей интересов (ROI) за относительно небольшое время выполнения. Как показал опыт использования программных решений (в частности ENVI), основным "узким" местом в производительности их алгоритмов являются операции ввода/вывода файловой подсистемы, а конкретнее — компромисс между использованием оперативной памяти и временем загрузки исходных данных в нее.

Согласно [4] исходное изображение для модуля атмосферной коррекции FLAASH должно иметь значение светимости для каждого пикселя в спектральном канале и быть в формате преобразования BIL (*band-interleaved-by-line*) или VIP (*band-interleaved-by-pixel*).

В формате BIL данные хранят значение для каждой линии снимка в порядке следования "канал за каналом", соответственно, для каждой строки (см. рис. 1, на четвертой стороне обложки). Формат VIP аналогичен BIL за исключением того, что данные в порядке следования "канал за каналом" хранятся для каждого пикселя (рис. 1) [2].

Поэтому, учитывая схему расположения данных в исходном изображении, запись их в файл назначения для конвертируемого формата, битность и тип данных, процесс конвертации форматов можно разбить на три этапа в целях оптимизации операции ввода/вывода, размещения и работы с массивами в оперативной памяти. Далее для краткости соответствующую конвертацию будем обозначать как BSQ  $\rightarrow$  BIL или BSQ  $\rightarrow$  VIP.

*Первый этап* аналогичен для BSQ  $\rightarrow$  BIL и BSQ  $\rightarrow$  VIP. Здесь осуществляется проверка общей длины

снимка в байтах в целях использования порционного считывания данных из файлов. Пороговое значение для порции установлено в 1 Гбайт. Выбрано оно опытным путем исходя из минимальных требований, необходимых для исполнения программного кода виртуальной машины Java, максимальной длины массива в Java и минимизации использования файла подкачки операционной системы Windows. На этом этапе вводятся следующие понятия: число линий снимка, размещенных в "heap size"  $^{heap}L$ , целое число таких размещений для общего числа линий  $^{num}L$  и остаток линий с учетом полного покрытия снимка  $^{remain}L$ :

$$^{heap}L = \frac{1024^3}{^{byte}L \cdot n}, \quad (1)$$

где  $^{heap}L \in \mathbb{Z}$  — число линий в "heap size" (в программном коде задано как тип Integer, поэтому результат деления — всегда целое число),  $^{byte}L \in \mathbb{Z}$  — длина одной линии в байтах,  $n \in \mathbb{Z}$  — числе спектральных каналов;

$$^{num}L = \frac{^{heap}L \cdot ^{heap}L}{^{num}L} + ^{remain}L, \quad (2)$$

где  $^{num}L \in \mathbb{Z}$  — число линий снимка для каждого спектрального канала,  $^{heap}L \in \mathbb{Z}$  — число размещений для общего числа линий,  $^{remain}L \in \mathbb{Z}$  — остаток линий с учетом полного покрытия снимка.

*Второй этап* также аналогичен для двух типов конвертаций BSQ  $\rightarrow$  BIL и BSQ  $\rightarrow$  VIP. На данном этапе выполняются самые длительные по времени операции. Здесь происходит порционное считывание данных из файлов спектральных каналов в массив типа BufferArray, содержащий данные  $i$ -й порции  $^{byte}A_k^i$  и остаток данных  $^{byte}A_k^{remain}$  для каждого из спектральных каналов:

$$^{byte}A_k^i = [_{offset}L + ^{heap}L \cdot ^{byte}L \cdot i, _{offset}L + ^{heap}L \cdot ^{byte}L \cdot (i + 1)]_k; \quad (3)$$

$$^{byte}A_k^{remain} = [_{offset}L + ^{heap}L \cdot ^{byte}L \cdot \frac{^{heap}L}{^{num}L}, _{offset}L + ^{heap}L \cdot ^{byte}L \cdot \frac{^{heap}L}{^{num}L} + ^{remain}L]_k, \quad (4)$$

где  $^{byte}A_k^i$  — элемент массива порции данных и  $^{byte}A_k^{remain}$  — остаток данных  $k$ -го спектрального канала;  $_{offset}L \in \mathbb{Z}$  — значение тега StripOffset TIFF-

файла,  $k = 0 \dots n - 1$ ,  $i = 0 \dots \frac{^{heap}L}{^{num}L} - 1$ ,  $k, i \in \mathbb{Z}$ . В формулах (3) и (4) в виде отрезков с границами показаны начальное и конечное смещения байтов в TIFF-файле  $k$ -го спектрального канала, соответствующие считываемой  $i$ -й порции данных или остатка данных.

Далее будем рассматривать  $^{byte}A_k^i$  и  $^{byte}A_k^{remain}$  как массивы длиной  $^{heap}L \cdot ^{byte}L$  и  $^{remain}L \cdot ^{byte}L$  соответственно.

На *третьем этапе* происходит размещение данных  $\text{byte}_{buffer}A$  в массив типа BufferArray согласно выбранному типу конвертации  $\text{BSQ} \rightarrow \text{BIL}$  [см. ниже формулы (7)–(10)] или  $\text{BSQ} \rightarrow \text{VIP}$  [см. ниже формулы (11)–(14)] с последующим сохранением его на диск. Здесь применяется мультипоточковый подход к одновременной записи данных  $\text{byte}_{buffer}A^i$  в выделенную память в соответствии со схемой расположения для  $k$ -го спектрального канала (см. рис. 1). Вводится переменная "число процессорных ядер", для конвертации  $\text{BSQ} \rightarrow \text{VIP}$  — переменная "число обрабатываемых спектральных каналов на ядро в порции данных"  $\text{core}B \in \mathbb{Z}$ , для конвертации  $\text{BSQ} \rightarrow \text{BIL}$  — переменная "разбивки порции данных по вычислительным ядрам"  $\text{core}(\text{heap}L) \in \mathbb{Z}$ :

$$\text{core}B = \frac{n}{\text{core}N}, n = \text{core}B \cdot \text{core}N + \text{core}_{\text{remain}}B, \quad (5)$$

где  $\text{core}N \in \mathbb{Z}$  — число процессорных ядер;  $\text{core}_{\text{remain}}B \in \mathbb{Z}$  — остаток обрабатываемых спектральных каналов на ядро в порции данных;

$$\text{core}(\text{heap}L) = \frac{\text{heap}L}{\text{core}N},$$

$$\text{heap}L = \text{core}(\text{heap}L) \cdot \text{core}B + \text{core}_{\text{remain}}(\text{heap}L), \quad (6)$$

где  $\text{core}_{\text{remain}}(\text{heap}L) \in \mathbb{Z}$  — остаток разбивки порции данных по вычислительным ядрам;

$$\text{BIL}^i[k \cdot \text{byte}L + j \cdot \text{byte}L \cdot n] = \lfloor \text{byte}_{buffer}A_k^i[\text{byte}L \cdot j],$$

$$\text{byte}_{buffer}A_k^i[\text{byte}L \cdot j + \text{byte}L], \quad (7)$$

где  $\text{BIL}^i$  — сконвертированная  $i$ -я ( $i = 0 \dots \frac{\text{heap}L}{\text{num}} - 1$ ) порция данных,  $j = c \cdot \text{core}(\text{heap}L) \dots c \cdot \text{core}(\text{heap}L) + \text{core}_{\text{remain}}(\text{heap}L)$ , где  $c = 0 \dots \frac{\text{heap}L - \text{core}_{\text{remain}}(\text{heap}L)}{\text{core}(\text{heap}L)}$ ,  $k = 0 \dots n - 1$ ,  $k, c, i, j \in \mathbb{Z}$ ;

$$\text{remainBIL}^i[k \cdot \text{byte}L + j \cdot \text{byte}L \cdot n] =$$

$$= \lfloor \text{byte}_{buffer}A_k^i[\text{byte}L \cdot j], \text{byte}_{buffer}A_k^i[\text{byte}L \cdot j + \text{byte}L], \quad (8)$$

где  $\text{remainBIL}^i$  — остаток сконвертированной  $i$ -й порции данных после многопоточковой обработки,  $j = \text{heap}L - \text{core}_{\text{remain}}(\text{heap}L) \dots \text{heap}L$ ,  $k = 0 \dots n - 1$ ;

$$\text{remainBIL}[k \cdot \text{byte}L + j \cdot \text{byte}L \cdot n] =$$

$$= \lfloor \text{byte}_{buffer}A_k^{\text{remain}}[\text{byte}L \cdot j], \text{byte}_{buffer}A_k^{\text{remain}}[\text{byte}L \cdot j + \text{byte}L], \quad (9)$$

где  $\text{remainBIL}$  — остаток данных после разбиения на порции согласно (2),  $j = 0 \dots \text{remain}L$ ,  $k = 0 \dots n - 1$ ;

$$\text{BIL} = \left( \bigcup_{i=0}^{\frac{\text{heap}L}{\text{num}} - 1} (\text{BIL}^i \cup \text{remainBIL}^i) \right) \cup \text{remainBIL}, \quad (10)$$

где  $\text{BIL}$  — полный файл гиперспектрального снимка с преобразованием  $\text{BIL}$ .

Соответственно, для конвертации  $\text{BSQ} \rightarrow \text{VIP}$   $i$ -й порции данных ( $i = 0 \dots \frac{\text{heap}L}{\text{num}} - 1$ ) применяли следующие формулы:

$$\text{VIP}^i[2j \cdot n + 2k] =$$

$$= \text{byte}_{buffer}A_k^i[2j], \text{VIP}^i[2j \cdot n + 2k + 1] =$$

$$= \text{byte}_{buffer}A_k^i[2j + 1], \quad (11)$$

где  $\text{VIP}^i$  — сконвертированная  $i$ -я порция данных,  $j = 0 \dots \frac{\text{heap}L \cdot \text{width}L}{\text{width}L}$ ;  $\text{width}L$  — ширина снимка, равная значению параметра PRODUCT\_SAMPLES в метаописании снимка,  $k = c \cdot \text{core}B \dots c \cdot \text{core}B + \text{core}B$ ,

$$c = 0 \dots \frac{n - \text{core}_{\text{remain}}B}{\text{core}B}, k, c, i, j \in \mathbb{Z};$$

$$\text{remainVIP}^i[2j \cdot n + 2k] = \text{byte}_{buffer}A_k^i[2j],$$

$$\text{remainVIP}^i[2j \cdot n + 2k + 1] = \text{byte}_{buffer}A_k^i[2j + 1], \quad (12)$$

где  $\text{remainVIP}^i$  — остаток сконвертированной  $i$ -й порции данных после многопоточковой обработки,  $j = 0 \dots \frac{\text{heap}L \cdot \text{width}L}{\text{width}L}$ ,  $k = n - \text{core}_{\text{remain}}B \dots n$ ,  $k, j \in \mathbb{Z}$ ;

$$\text{remainVIP}[2j \cdot n + 2k] = \text{byte}_{buffer}A_k^{\text{remain}}[2j],$$

$$\text{remainVIP}[2j \cdot n + 2k + 1] = \text{byte}_{buffer}A_k^{\text{remain}}[2j + 1], \quad (13)$$

где  $\text{remainVIP}$  — остаток данных после разбиения на порции согласно (2),  $j = 0 \dots \text{remain}L \cdot \text{width}L$ ,  $k = 0 \dots n$ ;

$$\text{VIP} = \left( \bigcup_{i=0}^{\frac{\text{heap}L}{\text{num}} - 1} (\text{VIP}^i \cup \text{remainVIP}^i) \right) \cup \text{remainVIP}, \quad (14)$$

где  $\text{VIP}$  — полный файл гиперспектрального снимка с преобразованием  $\text{BIL}$ .

Программная реализация алгоритма представлена как расширение для программного продукта Exelis ENVI (рис. 2, см. четвертую сторону обложки). Графический интерфейс пользователя разработан собственными средствами пакета на базе технологии GUI-WIDGETS в интеграции с пакетами Java SwingX. Для взаимодействия с Java-классами, реализующими логику представленного алгоритма, использовали технологию Java-Bridge IDL. Данная технология позволила инкапсулировать сложные процедуры расчетов от конечного пользователя и представить конечный интерфейс как обычное расширение ENVI IDL. Благодаря технологии Java-Bridge IDL разработанные программные компоненты Java могут быть использованы в сторонних программных продуктах, поддерживающих IDL. Гибкая настройка конфигурационных файлов Java-Bridge IDL позволила контролировать распределение машинных ресурсов на программном

уровне, за счет передачи параметров запуска виртуальной машины Java, "на лету".

Ниже представлены результаты тестирования алгоритма (рис. 3, см. четвертую сторону обложки). Тестирование проводили на пяти гиперспектральных снимках, полученных с ресурса [9]. Были выбраны снимки размерами 1960 × 6930, 242 спектральных каналов. Для каждого снимка проводили по 25 "проходов" алгоритмов (базового и многопоточного). На графиках рис. 3 (см. четвертую сторону обложки) представлены средние значения времени работы алгоритмов. Значения для двух типов конвертаций BSQ → VIP и BSQ → BIL оказались приблизительно одинаковы, отличие составило около ±5 с.

### Заключение

Применение мультипоточного подхода позволило существенно сократить время (в 10 раз) загрузки снимка и преобразований (BIL, VIP, BSQ). Этого удалось добиться путем разбиения процесса обработки гиперспектрального изображения по каналам, либо попиксельно, где в первом случае общее число каналов разделено равномерно по числу вычислительных ядер процессора(-ов), во втором случае разбиение идет по ширине и/или по высоте снимка, а также путем порционной загрузки данных снимка в оперативную память, что позволило существенно снизить нагрузку на файловую систему ввода/вывода и привело к снижению времени выполнения данных операций.

Использование технологии IDL-Java Bridge дало возможность интеграции разработанных программных компонентов в среде ENVI IDL. Предложенный алгоритм способен эффективно выполняться на платформах с низкопроизводительной файло-

вой системой и емкостью оперативной памяти (до 2 Гбайт).

Предложенный подход может быть реализован на вычислительных кластерах, что резко уменьшает время пакетной обработки гиперспектральных снимков.

### Список литературы

1. **Adler-Golden S. M., Perkins T., Matthew M. W., Berk A., Bernstein L. S.**, et al. Speed and accuracy improvements in FLAASH atmospheric correction of hyperspectral imagery // SPIE Optical Engineering. 2012. Vol. 51 (11). P. 111707 (1–10).
2. **BIL, VIP, and BSQ raster files.** // ESRI. ArcGIS 9.2 Desktop Help. URL: [http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=BIL,\\_VIP,\\_and\\_B\\_SQ\\_raster\\_files](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=BIL,_VIP,_and_B_SQ_raster_files) (дата обращения 22.07.2014).
3. **EarthExplorer.** // USGS. URL: <http://earthexplorer.usgs.gov/> (дата обращения 22.07.2014)
4. **Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH)** // Exelis ENVI. URL: <http://www.exelisvis.com/docs/FLAASH.html> (дата обращения 24.07.2014).
5. **Hyperion level 1gst (L1GST) product output files data format control book.** Earth Observing-1 (EO-1). Version 1.0. Department of the Interior U.S. Geological Survey. 2006. 24 p.
6. **Perkins T., Adler-Golden S. M., Cappelaere P., Mandl D.** High-speed Atmospheric Correction for Spectral Image Processing // SPIE Proceeding: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII. 2012. Vol. 8390. P. 245–252.
7. **Qu Z., Goetz A. F. H., Kindel B.** High-accuracy atmospheric correction for hyperspectral data (HATCH) model // Geoscience and Remote Sensing. 2003. Vol. 41 (6). P. 1223–1231.
8. **San B. T., Suzen M. L.** Evaluation of different atmospheric correction algorithms for EO-1 Hyperion imagery // International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science. Tokyo. 2010. Vol. 38 (8). P. 392–397.
9. **The IDL Thread Pool.** // Exelis ENVI. URL: [http://www.exelisvis.com/docs/TheThread\\_Pool.html](http://www.exelisvis.com/docs/TheThread_Pool.html) (дата обращения 29.09.2014).
10. **Thompson B. J., Rahman Z., Park S. K.** Multiscale retinex for improved performance in multispectral image classification // SPIE Proceedings: Visual Information Processing IX. 2000. Vol. 4041. P. 34–44.

**V. P. Potapov**, Deputy Managing Director, Branch Manager, potapov@ict.sbras.ru,  
**S. E. Popov**, Chief Scientist, popov@ict.sbras.ru  
Institute of Computational Technologies SB RAS, Novosibirsk, 630090, Russia

## Multithreaded Algorithm for Band Interleave Conversion of Hyperspectral Images

*This paper presents a high-performance data conversion algorithm of spectral bands (Band Interleave Conversion) of the sensor EO-1 Hyperion, providing the ability to run it on multiprocessor platforms in multi-threaded mode and ensures the effective implementation on low I/O systems. Proposed implementation of proportional reading image data into RAM, followed by the placement of values (Digitals Number) in the target buffer arrays in multiple threads, calculated by the number of spectral channels or the number of lines of the image. Results of testing the algorithm presented with their basic counterparts software package Exelis ENVI. It is shown that the speed (second) of the multithreaded algorithm hundreds of times higher than for the native algorithm. In particular, the process of multithreaded band interleave conversion on the 8-core architecture with the use RAMDisk I/O-subsystem was only 45 seconds, while the native algorithms needed in similar environment of about 6800 seconds. The paper presents an Exelis ENVI extension, realizing the developed algorithm based on the GUI-WIDGETS integration with packages Java SwingX using IDL-Bridge technology.*

**Keywords:** converting the order of the data, spectral channels, hyperspectral images, multithreading, java, IDL-Bridge, EO-1 Hyperion

## References

1. **Adler-Golden S. M., Perkins T., Matthew M. W., Berk A., Bernstein L. S., et al.** Speed and accuracy improvements in FLAASH atmospheric correction of hyperspectral imagery, *SPIE Optical Engineering*, 2012, Vol. 51 (11), pp. 111707 (1–10).
2. **BIL, BIP, and BSQ** raster files. *ESRI. ArcGIS 9.2 Desktop Help*. Retrieved from. URL: [http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=BIL,\\_BIP,\\_and\\_BSQ\\_raster\\_files](http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=BIL,_BIP,_and_BSQ_raster_files)
3. **EarthExplorer**. *USGS*. Retrieved from <http://earthexplorer.usgs.gov/>.
4. **Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH)**, *Exelis ENVI*. Retrieved from <http://www.exelisvis.com/docs/FLAASH.html>
5. **Hyperion** level 1 gst (L1GST) product output files data format control book. *Earth Observing-1 (EO-1)*. Version 1.0. Department of the Interior U. S. Geological Survey, 2006, 24 p.
6. **Perkins T., Adler-Golden S. M., Cappelaere P., Mandl D.** High-speed Atmospheric Correction for Spectral Image Processing, *SPIE Proceeding: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII*, 2012, Vol. 8390, pp. 245–252.
7. **Qu Z., Goetz A. F. H., Kindel B.** High-accuracy atmospheric correction for hyperspectral data (HATCH) model, *Geoscience and Remote Sensing*, 2003, Vol. 41 (6), pp. 1223–1231.
8. **San B. T., Suzen M. L.** Evaluation of different atmospheric correction algorithms for EO-1 Hyperion imagery, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, Tokyo, 2010, vol. 38 (8), pp. 392–397.
9. **The IDL Thread Pool**. *Exelis ENVI*. Retrieved from [http://www.exelisvis.com/docs/The\\_Thread\\_Pool.html](http://www.exelisvis.com/docs/The_Thread_Pool.html).
10. **Thompson B. J., Rahman Z., Park S. K.** Multiscale retinex for improved performance in multispectral image classification, *SPIE Proceedings: Visual Information Processing IX*, 2000, vol. 4041, pp. 34–44.

УДК 621.391

**С. В. Дворников**, д-р техн. наук, проф., профессор каф.,  
**А. В. Пшеничников**, канд. техн. наук, доц., доц. каф.,  
**С. С. Манаенко**, канд. техн. наук, ст. преподаватель каф.,  
Военная академия связи, г. Санкт-Петербург, e-mail: manaenkoss@mail.ru

## Помехоустойчивая модель сигнала КАМ-16 с трансформированным созвездием

*Представлены результаты исследования, направленные на разработку модели сигнала КАМ-16 с трансформированной конstellационной диаграммой. Определяются показатели помехоустойчивости, по которым оценивается разработанная модель и модель с трансформированным созвездием в соответствии со стандартом цифрового телевидения DVB-T2.*

**Ключевые слова:** помехоустойчивость, конstellационная диаграмма, вектор сигнального созвездия, квадратурная амплитудная манипуляция

### Введение

В настоящее время в радиотехнике широко используются сигнальные конструкции (СК), формируемые квадратурным методом на основе синфазной ( $I$ ) и квадратурной ( $Q$ ) составляющих сигнала. В частности, к таковым относятся сигналы квадратурной амплитудной манипуляции (КАМ). Каждый сигнал КАМ можно изобразить в виде сигнальной точки, координаты которой определяются значениями  $Q$  и  $I$ . Совокупность таких точек образует так называемое сигнальное созвездие (конstellационную диаграмму). При этом наибольшее распространение в радиотехнике получили созвездия, имеющие прямоугольную диаграмму из 16 сигнальных точек (КАМ-16) [1, 2].

Как правило [3], помехоустойчивость СК определяется минимальным евклидовым расстоянием (МЕР)  $d_E$ . Указанный показатель представляет собой наименьшее геометрическое расстояние между точками сигнального созвездия. Другим важным

энергетическим показателем СК является отношение максимальной мощности СК к ее среднему значению, определяемому в работе [4] как пик-фактор (ПФ). По показателю МЕР сигналы КАМ являются более помехоустойчивыми по отношению к сигналам фазовой манипуляции такой же размерности (позиционности), однако по показателю ПФ они проигрывают последним.

Между тем в телевизионном стандарте DVB-T2 для повышения помехоустойчивости рекомендуется использовать модели СК с трансформированными конstellационными диаграммами, полученными путем поворота всех точек сигнального созвездия на одинаковый угол. В результате применения указанной процедуры каждая точка сигнального созвездия получает независимые координаты, которые дают возможность на приемном конце восстановить позиции точки сигнального вектора конstellационной диаграммы СК даже по одной координате, и тем самым повысить помехоустойчивость процедур демодуляции.

Следует отметить, что рекомендованная в стандарте DVB-T2 трансформация не изменяет энергетические параметры модели СК, т. е. повышение помехоустойчивости достигается только за счет получения независимых координат каждой из точек сигнального созвездия.

Новизна работы заключается в разработке подхода оценки помехоустойчивости СК на основе их энергетических параметров, в том числе и значения МЕР. В связи с этим в статье представлена модель СК, в которой повышение помехоустойчивости достигается прежде всего за счет изменения ее энергетических параметров.

### Энергетические параметры сигнальных конструкций

В общем случае математическую модель канала связи для произвольного сигнала  $s(t)$  можно представить выражением

$$z(t) = \mu(t)s(t) + \xi(t), \quad 0 \leq t \leq T, \quad (1)$$

где  $\mu(t)$  — коэффициент передачи канала;  $\xi(t)$  — аддитивный белый Гауссов шум (АБГШ) с односторонней спектральной плотностью шума  $N_0$ ;  $T$  — длительность символа.

Заметим, что при  $\mu(t) = 1$  в выражении (1) получаем классический канал с АБГШ. Если же  $\mu(t) = \mu$ , где  $\mu$  — случайная величина, то получаем модель канала с неселективными замираниями. При условии  $T < \tau_s$ , где  $\tau_s$  — интервал замираний, фазу принимаемого сигнала  $z(t)$  можно оценить без ошибок.

В соответствии с теоремой ортогонализации Грамма—Шмидта [3] произвольный сигнал  $s(t)$  может быть представлен в виде

$$s(t) = \sum_{v=1}^N s_v w(t),$$

где  $w(t)$  — базисные функции.

Такой подход позволяет перейти к геометрической интерпретации сигналов и рассматривать их в конечном евклидовом пространстве. При этом энергия любого сигнала может быть определена как

$$E = \sum_{v=1}^N s_v^2.$$

Поскольку энергия конечна,  $E < \infty$ , то очевидно, что среди всех возможных моделей СК есть конструкции, имеющие максимальную энергию, которую будем обозначать  $E_m$ .

Тогда средняя энергия сигнала будет определяться как

$$E_c = \sum_{r=0}^{M-1} p(r) E_r,$$

где  $p(r)$  — априорная вероятность передачи  $r$ -й составляющей СК.

Для условия равновероятных состояний всех составляющих СК, т. е. когда  $p(r) = 1/M$ , где  $M$  — позиционность сигнала, имеем [5]

$$E_c = \frac{1}{M} \sum_{r=0}^{M-1} E_r.$$

Рассмотрим энергетические показатели модели сигнала  $s(t)$ , одним из которых является ПФ, рассчитываемый как отношение максимальной энергии к ее среднему значению:

$$\Pi^2 = E_m/E_c. \quad (2)$$

Другим важным показателем, характеризующим энергетические свойства СК, является МЕР, определяющий минимальную геометрическую разность между двумя сигналами  $d_{rk} = \|s_r - s_k\|$ , или

$$d_{rk} = \sqrt{\sum_{v=1}^N (s_{rv} - s_{kv})^2}.$$

В этом случае МЕР можно определить как

$$d_E = \min_{r \neq k} d_{r,k}, \quad r, k = \overline{0, M-1},$$

отличную от нуля, если  $r \neq k$ .

В своей совокупности энергетические показатели определяют коэффициент помехоустойчивости СК:

$$g_m = \left( \frac{d}{2\sqrt{E_m}} \right)^2 \quad \text{и} \quad g_c = \left( \frac{d}{2\sqrt{E_c}} \right)^2.$$

Отношение максимальной энергии и средней энергии к односторонней спектральной плотности шума определяется, соответственно, как

$$h_m^2 = \frac{E_m}{N_0} \quad \text{и} \quad h_c^2 = \frac{E_c}{N_0}.$$

Тогда с учетом полученных формул значение максимальной энергии сигнала можно представить следующим образом [6]:

$$h_m^2 = \Pi_c^2 h_c^2 \frac{E_m}{N_0}.$$

При таком подходе величины  $E_{bm} = \frac{E_m}{\log_2 M}$  и

$E_{bc} = \frac{E_c}{\log_2 M}$  — соответственно, представляют собой максимальное и среднее значения энергии, затраченной для передачи одного бита, а  $h_{bm}^2 = \frac{E_{bm}}{N_0}$ ,

$$h_{bc}^2 = \frac{E_{bc}}{N_0}, \quad h_{bm}^2 = \Pi_c^2 h_{bc}^2 \quad \text{и} \quad h_c^2 = h_{bc}^2 \log_2 M.$$

Таким образом, поскольку помехоустойчивость любой модели СК определяется только ее энергетическими показателями, то именно их целесообразно использовать в качестве критерия оценки для сравнения моделей СК между собой. Причем равенство энергетических показателей различных моделей СК дает возможность утверждать, что они обладают одинаковыми свойствами помехоустойчивости.

### Оценка помехоустойчивости модели сигнала КАМ-16 с трансформированным созвездием

Для повышения помехоустойчивости в телевизионном стандарте DVB-T2 сигналы КАМ с трансформированными конstellационными диаграммами передаются по каналу на основе технологии (OFDM) (*orthofonal frequency division multiplexing*) [7].

Сама же трансформация по стандарту DVB-T2 заключается в повороте всех точек сигнального созвездия на  $16,8^\circ$  (рис. 1) [5]. В результате такого поворота каждая точка конstellационной диаграммы получает свои уникальные координаты по осям  $Q$  и  $I$ , которые принадлежат только ей, т. е. значения  $I_A$  и  $Q_A$  больше не встречаются ни у одной из точек сигнального созвездия, кроме точки  $A$  (рис. 1).

Указанный подход позволяет восстановить исходное положение точки  $A$  всего по одному из значений  $I_A$  или  $Q_A$ .

Между тем рассматриваемая трансформация не является единственно возможной. В частности, предлагается другой подход, реализация которого позволяет получить структурную модель созвездия, представленную на рис. 2. Аналитическая мо-

дель сигнала КАМ-16 с трансформированным созвездием имеет вид

$$s_r(t) = I_r \cos(2\pi f_0 t + \varphi_r) + Q_r \sin(2\pi f_0 t + \varphi_r),$$

где  $\varphi_r$  — угол поворота сигнальной точки;  $f_0$  — несущая частота;  $r = 1, 2, \dots, M$ .

В предлагаемой модели СК для векторов сигнальных созвездий каждого из квадрантов конstellационной диаграммы реализовано свое независимое вращение на  $17,1^\circ$ . В качестве условий трансформации рассматривались сохранение значения МЕР и обеспечение для каждой точки векторов сигнального созвездия пары ее уникальных координат по осям  $I$  и  $Q$ .

Заметим, что в относительных значениях координат, представленных на рис. 1 и 2, значения МЕР одинаковы у обоих созвездий, что дает основание предположить, что по данному показателю помехоустойчивости рассматриваемые модели близки друг к другу.

Однако расчет средней энергии и ПФ показал, что значения совокупности этих показателей предпочтительней у модели сигнала КАМ-16, созвездие которой трансформировано на основе предлагаемого подхода.

В частности, согласно проведенным расчетам по показателю ПФ, выигрыш составит около 8 %. При этом проигрыш в средней энергии составит не более 1,2 %. Следовательно, можно утверждать о возможном энергетическом выигрыше в помехоустойчивости, полученном за счет предложенной трансформации сигнального созвездия КАМ-16.

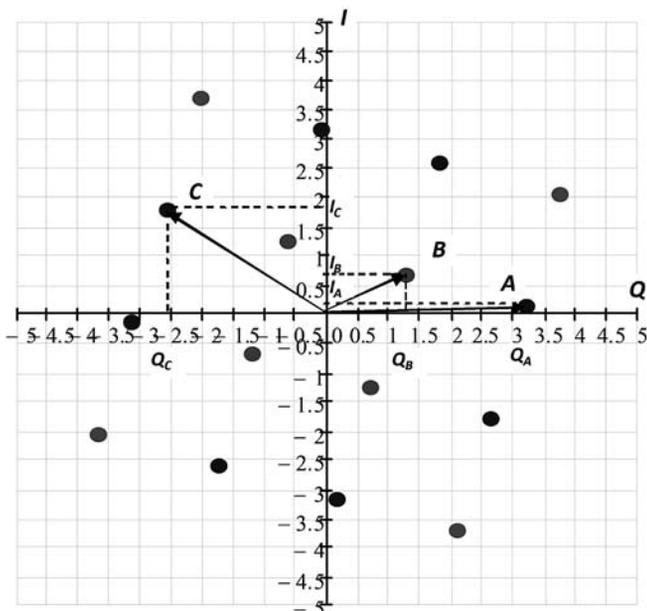


Рис. 1. Конstellационная диаграмма сигнала КАМ-16, трансформированная по стандарту DVB-T2

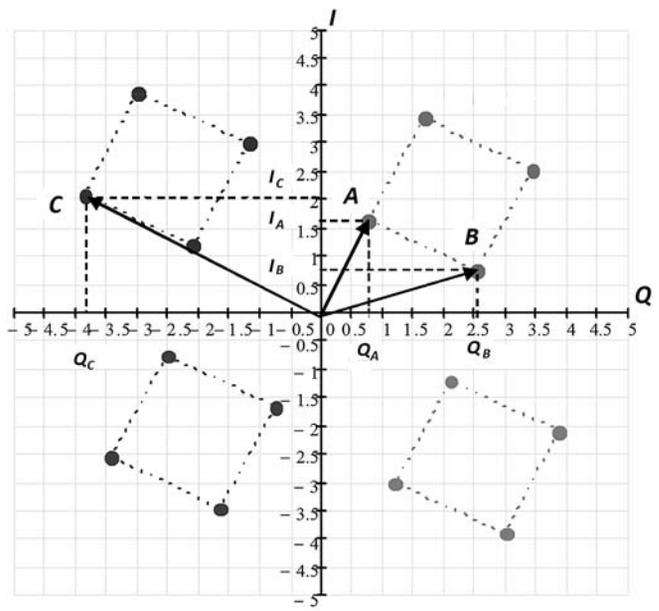


Рис. 2. Конstellационная диаграмма модели сигнала КАМ-16, трансформированная по альтернативному варианту

## Результаты сравнительной оценки помехоустойчивости моделей сигнальных конструкций КАМ-16

В целях проверки полученных результатов было проведено компьютерное моделирование процесса приема элементов модели сигнала КАМ-16 по стандарту DVB-T2, а также модели сигнала КАМ-16 с трансформированным созвездием. Результаты моделирования представлены на рис. 3.

Помехоустойчивость приема сигналов КАМ-16 определяли с позиций ошибочного приема одного бита информации. В качестве анализируемого показателя была выбрана вероятность ошибки на бит, характеризующая частоту проявления одиночных ошибок в канале.

Указанный показатель, определяющий зависимость вероятности ошибки на бит ( $P_{bc}$ ) от энергетических показателей СК, представлен в работе [6]:

$$P_{bc} = \frac{2}{\sqrt{M}} \log_2 \sqrt{M} \times \sum_{k=1}^{\log_2 \sqrt{M}} \sum_{i=0}^{(1-2^{-k})\sqrt{M}-1} \left( (-1)^{\frac{i2^{k-1}}{\sqrt{M}}} \left( 2^{k-1} - \left( \frac{i2^{k-1}}{\sqrt{M}} + \frac{1}{2} \right) \right) \right) \times Q \left( (2i+1) \sqrt{\frac{6 \log_2 \sqrt{M} E_{bc}}{2(M-1) N_0}} \right), \quad (3)$$

где  $M$  — позиционность СК;  $\frac{E_{bc}}{N_0}$  — отношение значения средней энергии бита к значению спектральной мощности шума.

По результатам анализа можно сделать следующие выводы.

Поворот всех векторов сигнального созвездия модели КАМ-16 на одинаковый угол (см. рис. 1) не изменяет ее помехоустойчивость в канале с АБГШ.

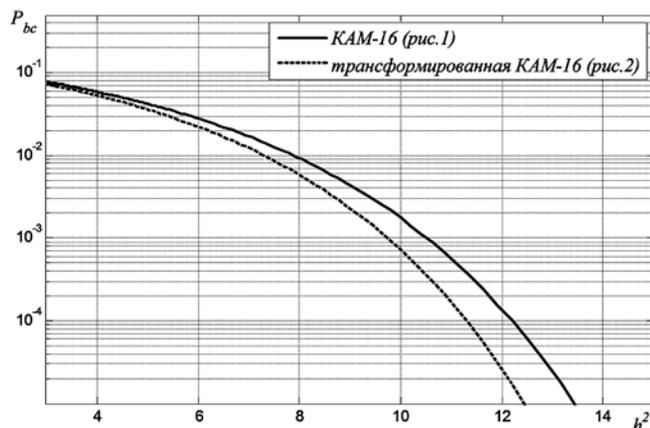


Рис. 3. Зависимость вероятности ошибки на бит от спектральной плотности мощности шума

Поворот векторов сигнального созвездия различных квадрантов констелляционной диаграммы на различные углы (см. рис. 2) позволяет улучшить энергетические параметры модели КАМ-16 по показателю помехоустойчивости. В результате, при приеме сигналов на основе такой модели снижается частота проявления одиночных ошибок при воздействии АБГШ.

## Заключение

Таким образом, предложенный подход к формированию модели сигнала КАМ-16 с трансформированной констелляционной диаграммой позволит повысить помехоустойчивость его приема по отношению к модели, используемой в стандарте DVB-T2, и получить сигнальное созвездие, каждая точка которого имеет уникальные координаты в пространстве  $I$  и  $Q$ . Причем разность градаций координатных значений по осям будет такой же, как и у модели созвездия, трансформированного по стандарту DVB-T2. Следует заметить, что предложенная модель СК может быть демодулирована и в соответствии со стандартом DVB-T2, однако при этом уровень помехоустойчивости приема снизится не менее чем на 6 дБ.

Повысить помехоустойчивость, по мнению авторов, позволит и совместная частотно-временная обработка сигналов в трактах приема [8].

Дальнейшее исследование авторы связывают с разработкой эффективного способа формирования модели сигнальной конструкции в соответствии с предложенным подходом и с проведением экспериментов, связанных с оценкой помехоустойчивости ее приема при различных способах кодирования точек сигнального созвездия.

## Список литературы

1. Smith J. G. Odd-bit quadrature amplitude-shift keying // IEEE Trans. Communications. March 1975. Vol. COM-23. P. 385—389.
2. Shi Q. Asymptotic clipping noise distribution and its impact on M-ary QAM transmission over optical fiber // IEEE Trans. Communications. June 1995. Vol. COM-43. P. 2077—2084.
3. Коржик В. И., Финк Л. М., Щелкунов К. Н. Расчет помехоустойчивости систем передачи дискретных сообщений: справочник. М.: Радио и связь, 1981. 232 с.
4. Кловский Д. Д. Передача дискретных сообщений по радиоканалам. М.: Радио и связь, 1982. 304 с.
5. Савищенко Н. В. Многомерные сигнальные конструкции: их частотная эффективность и помехоустойчивость приема / Под ред. Д. Л. Бураченко. СПб.: Изд-во Политехн. ун-та, 2005. 420 с.
6. Прокис Дж. Цифровая связь / Пер. с англ. под ред. Д. Д. Кловского. М.: Радио и связь, 2000. 800 с.
7. Framing structure, channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). DVB Document A122, Jun. 2008.
8. Алексеев А. А., Аладинский В. А., Дворников С. В., Железняк В. К., Комарович В. Ф. Применение методов частотно-временной обработки акустических сигналов для анализа параметров реверберации // Научное приборостроение. 2001. Т. 11, № 1. С. 65—76.

## Increased Noise Immunity Signal 16-QAM Constellation with Transformed

Currently, the television standard DVB-T2 to improve noise immunity, it is recommended to use a model of the signal of a quadrature manipulation transformed constellations diagrams obtained by turning all points of the signal constellation at the same angle. As a result of applying this procedure, each point of the signal constellation receives independent coordinates, which give the opportunity for the receiving end to restore the position of the signal vector constellation charts even for one coordinate, and thereby improve the robustness of the procedures demodulation. When this recommended standard DVB-T2 transformation does not change the energy parameters of the model signal design, and increased robustness is achieved only by obtaining independent of the coordinates of each point of the signal constellation.

In the paper, the rotation vectors of the signal constellation various quadrants constellational chart to make different angles, which will improve the energy parameters of the signal model QAM-16 in terms of noise immunity. Presents the results of a study aimed at developing a model for such a signal, determines the noise immunity by which to evaluate the developed model and the model with transformed constellation in accordance with the standard DVB-T2.

**Keywords:** immunity, constellation diagram, vector signal constellation, quadrature amplitude modulation

### References

1. **Smith J. G.** Odd-bit quadrature amplitude-shift keying, *IEEE Trans. Communications*, March 1975, vol. COM-23, pp. 385–389.
2. **Shi Q.** Asymptotic clipping noise distribution and its impact on M-ary QAM transmission over optical fiber, *IEEE Trans. Communications*, June 1995, vol. COM-43, pp. 2077–2084.
3. **Korzhih V. I., Fink L. M., Shhelkunov K. N.** *Raschet pomehoustojchivosti sistem peredachi diskretnyh soobshhenij: spravochnik.* Moscow, Radio i svjaz', 1981. 232 p.
4. **Klovskij D. D.** *Peredacha diskretnyh soobshhenij po radiokanalam.* Moscow, Radio i svjaz', 1982. 304 p.
5. **Savishhenko N. V.** *Mnogomernye signal'nye konstrukcii: ih chastotnaja jeffektivnost' i pomehoustojchivost' priema*, Ed. D. L. Bura-chenko. SPb.: Izdatel'stvo Politehnicheskogo universiteta, 2005. 420 p.
6. **Prokis Dzh.** *Cifrovaja svjaz'*, per. s angl, ed. D. D. Klovskogo. Moscow: Radio i svjaz', 2000, 800 p.
7. **Framing** structure, channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). *DVB Document A122*, Jun. 2008.
8. **Alekseev A. A., Aladinskij V. A., Dvornikov S. V., Zhelez-njak V. K., Komarovich V. F.** *Primenenie metodov chastotno-vremennoj obrabotki akusticheskikh signalov dlja analiza parametrov reverberacii.* *Nauchnoe priborostroenie*, 2001, vol. 11, no. 1, pp. 65–76.

УДК 519.688, 669.14.242

**Н. Н. Светушков**, канд. техн. наук, доц. каф., e-mail: svt.n.n@mail.ru

Московский авиационный институт (национальный исследовательский университет), Москва, Россия

## Параллельный метод струн для численного решения нелинейных задач теплопроводности

Представлен способ организации параллельных вычислительных потоков для решения нелинейных задач нестационарной теплопроводности на основе разработанного ранее метода струн. В методе струн использован новый алгоритм, при котором уравнения в частных производных сведены к эквивалентной системе интегральных уравнений. Приведена структурная схема вычислительных потоков и показано, что наибольшее количество процессов может быть организовано в случае, когда в каждом потоке решается одно интегральное уравнение. Интегральное описание, лежащее в основе параллельного подхода, позволяет не только более эффективно моделировать процессы теплопередачи, подавляя осцилляционный характер поведения решения в случае больших градиентов температур, но и естественным образом учитывать меняющиеся с течением времени теплофизические характеристики изделия. Проведены тестовые расчеты для модельной задачи и показано, что значительная эффективность алгоритма может быть достигнута для задач большой размерности и в случае использования многопроцессорной техники (включая графические платы). Разработанный алгоритм может быть применен для расчета температурных полей в микровелах, при термической обработке заготовок из жаропрочных сталей, а также при формовании изделий из препрегов для авиационной и ракетной промышленности.

**Ключевые слова:** численные методы, параллельные вычисления, программные средства, моделирование, уравнения теплопроводности, интегральные уравнения, точность вычислений

## Введение

Широкое применение вычислительной техники при решении практически важных задач часто сдерживается сложностью применяемых алгоритмов и значительными временными затратами. В частности, это относится к моделированию разнообразных технологических процессов, включая такие, как термическая обработка металлов и сплавов, которые используются в металлургии, а также при нанесении жаропрочных покрытий в авиационной промышленности. Значительно уменьшить вычислительные затраты при численном решении задач моделирования позволяет применение как многопроцессорных комплексов, так и разработка специализированных многопоточных вычислительных процедур, организованных на графических ускорителях (графических платах). Несмотря на то что параллельным вычислениям посвящено большое число работ [например 1—6] и известна общая структура используемых программных средств, часто для каждой конкретной задачи необходимо применять собственный уникальный прием. Если рассматривать уравнения в частных производных, то применение простых явных схем позволяет достаточно эффективно проводить распараллеливание вычислительного процесса [2] (геометрический параллелизм), однако при больших временных промежутках здесь стоит вопрос о точности получаемых результатов, который связан со сходимостью прямых схем (в одномерном случае — аналог метода Эйлера). Если же задача является нелинейной, например, когда характеристики среды зависят от искомых величин (в данном случае температуры), то возникает необходимость организовывать различные итерационные процедуры. Применение явных схем в этом случае практически невозможно или приводит к слишком большим погрешностям. Таким образом, решение нелинейных (по теплофизическим характеристикам) задач теплопроводности приводит к необходимости использовать различные неявные схемы, которые, как известно [1—4], плохо поддаются распараллеливанию.

Эффективность распараллеливания обусловлена не только возможностью организации большого числа независимых вычислительных потоков, но и возможностью каждого из потоков наиболее равномерно загрузить вычислительные мощности — такое свойство называется сбалансированностью параллельного алгоритма. Другими словами, организовывать вычисления необходимо таким образом, чтобы каждый процесс занимал по возможности одинаковый промежуток вычислительного времени, так как очевидно, что время выполнения всей задачи не может быть меньше, чем время выполнения самой трудоемкой из подзадач. Довольно часто выполнение этого условия зависит не только от применяемых численных алгоритмов, но и особенностей конкретной вычислительной задачи [5, 6].

Кроме этого, использование слишком большого числа вычислительных потоков упирается в про-

блему синхронизации и буферизации данных, при которой передача информации может потребовать сопоставимых временных затрат по сравнению с решением отдельной подзадачи. Таким образом, желательно, чтобы применяемые алгоритмы обеспечивали достаточную (инженерную) численную точность при ограниченном количестве вычислительных процедур.

Все эти проблемы остро стоят и при попытке смоделировать процессы распространения теплоты в пространственно протяженных и геометрически сложных неоднородных объектах, теплофизические характеристики которых зависят от температуры (к этому классу относится, например, и классическая задача Стефана). Не случайно новые подходы для решения такого рода задач вызывают пристальный интерес. Например, в Национальном Центре применения суперкомпьютерных вычислений (США, университет Иллинойса <http://www.ncsa.illinois.edu/>) специально выделены вычислительные ресурсы под проект "Прямое численное моделирование процессов теплопередачи в среде жидкость — твердое тело на основе многочастичной модели" ("*Modeling fluid-solid heat transfer using particle-resolved direct numerical simulation*", Shankar Subramaniam, Iowa State University).

Как известно [7, 8], в основном численные алгоритмы решения задач теплопроводности ограничиваются применением различных модификаций метода конечных элементов (МКЭ) и методами конечных разностей (МКР). Применение этих методов при решении нестационарной задачи теплопроводности на мелкомасштабном сеточном разбиении для пространственно протяженных трехмерных объектов приводит к большим временным затратам. К таким задачам можно отнести, например, процессы тепловыделения в ядерных реакторах и микротрещинах при их эксплуатации, а также термическую обработку заготовок из жаропрочных и жаростойких сплавов [10] или процессы формования нагружаемых изделий (фюзеляжа и крыльев) из препрегов в авиационной промышленности. В данной работе рассматривается задача значительного уменьшения вычислительного времени при решении такого рода задач путем организации сбалансированной многопоточковой вычислительной процедуры на основе интегрального "метода струн" [9].

### Организация параллельных вычислений в методе струн

Дифференциальные уравнения теплопроводности для неоднородных сред с изменяющимися теплофизическими характеристиками, как известно, имеют следующий вид (двумерный случай):

$$\begin{aligned} c\rho(u) \frac{\partial u}{\partial t} &= \frac{\partial}{\partial x} k_x(u) \frac{\partial u}{\partial x} + \frac{\partial}{\partial y} k_y(u) \frac{\partial u}{\partial y} + q; \\ u(x, y, 0) &= f(x, y); (x, y) \in \Omega; \\ W|_{(x, y) \in \Gamma(\Omega)} &= W_\Gamma, \end{aligned} \quad (1)$$

где  $u = u(x, y, t)$  — распределение температурного поля;  $f(x, y)$  — начальное распределение температур;  $c\rho(u)$  — произведение удельной теплоемкости и плотности материала, зависящие как от температуры, так и от пространственных координат;  $k_{x, y}(u)$  — коэффициенты теплопроводности по соответствующим осям, также зависящие от температуры;  $\Omega$  — область решения задачи;  $\Gamma(\Omega)$  — граничная поверхность;  $W_\Gamma = W_\Gamma(t)$  — зависящие от времени плотности тепловых потоков на границе области;  $q = q(x, y, t)$  — функция плотности распределения тепловых источников.

Искомые функции в интегральных уравнениях метода струн [9] являются с точностью до аддитивных функций производные от тепловых потоков по соответствующим направлениям (функции  $S_x$  и  $S_y$ ):

$$\begin{aligned} & \int_0^t \frac{1}{c\rho(u)} S_x(x, y, \tau) d\tau + \int_{x^*(y)}^x \frac{d\xi}{k_x(u)} \left[ \bar{r}_x(\xi, y, t) \times \right. \\ & \times \left. \int_{\underline{x}(y)}^{\xi} S_x(\xi_1, y, t) d\xi_1 - \underline{r}_x(\xi, y, t) \int_{\xi}^{\bar{x}(y)} S_x(\xi_1, y, t) d\xi_1 \right] = \\ & = -\underline{w}_x(y, t) \int_{x^*(y)}^x \frac{\bar{r}_x(\xi, y, t)}{k_x(u)} d\xi - \bar{w}_x(y, t) \int_{x^*(y)}^x \frac{\underline{r}_x(\xi, y, t)}{k_x(u)} d\xi - \\ & - u_0(x, y) - \int_0^t \frac{q(x, y, \tau)}{c\rho(u)} d\tau - \int_0^t \frac{S_y(x, y, \tau)}{c\rho(u)} d\tau - \\ & - \int_0^t \left[ \int_{\underline{y}(x)}^{\bar{y}(x)} c\rho(u) d\eta \right]^{-1} \times \\ & \times \left[ - \int_{\underline{y}(x)}^{\bar{y}(x)} S_y(x, \eta, \tau) d\eta - \bar{w}_y(\tau) + \underline{w}_y(\tau) \right] dt; \quad (2) \end{aligned}$$

$$\begin{aligned} & \int_0^t \frac{S_y(x, y, \tau)}{c\rho(u)} d\tau + \int_{y^*(x)}^y \frac{d\eta}{k_y(u)} \left[ \bar{r}_y(x, \eta, t) \int_{\underline{y}(x)}^{\eta} S_y(x, \eta_1, t) d\eta_1 - \right. \\ & - \underline{r}_y(x, \eta, t) \int_{\eta}^{\bar{y}(x)} S_y(x, \eta_1, t) d\eta_1 \left. \right] = \\ & = -\underline{w}_y(x, t) \int_{y^*(x)}^y \frac{\bar{r}_y(x, \eta, t)}{k_y(u)} d\eta - \bar{w}_y(x, t) \int_{y^*(x)}^y \frac{\underline{r}_y(x, \eta, t)}{k_y(u)} d\eta - \\ & - u_0(x, y) - \int_0^t \frac{q(x, y, \tau)}{c\rho(u)} d\tau - \int_0^t \frac{S_x(x, y, \tau)}{c\rho(u)} d\tau - \\ & - \int_0^t \left[ \int_{\underline{x}(y)}^{\bar{x}(y)} c\rho(u) d\xi \right]^{-1} \times \\ & \times \left[ - \int_{\underline{x}(y)}^{\bar{x}(y)} S_x(\xi, y, \tau) d\xi - \bar{w}_x(\tau) + \underline{w}_x(\tau) \right] dt, \quad (3) \end{aligned}$$

где

$$\left\{ \begin{aligned} \underline{r}_x(x, y, t) &= \int_{\underline{x}(y)}^x c\rho(u) d\xi \Big/ \int_{\underline{x}(y)}^{\bar{x}(y)} c\rho(u) d\xi; \\ \bar{r}_x(x, y, t) &= \int_{\bar{x}(y)}^x c\rho(u) d\xi \Big/ \int_{\underline{x}(y)}^{\bar{x}(y)} c\rho(u) d\xi; \\ \underline{r}_y(x, y, t) &= \int_{\underline{y}(x)}^y c\rho(u) d\xi \Big/ \int_{\underline{y}(x)}^{\bar{y}(x)} c\rho(u) d\xi; \\ \bar{r}_y(x, y, t) &= \int_{\bar{y}(x)}^y c\rho(u) d\xi \Big/ \int_{\underline{y}(x)}^{\bar{y}(x)} c\rho(u) d\xi. \end{aligned} \right. \quad (4)$$

Если решение системы уравнений (2)—(3) найдено, то температура может быть вычислена путем простого интегрирования искомых функций:

$$\begin{aligned} u(x, y, t) &= u_0(x, y) + \int_0^t \frac{q(x, y, \tau)}{c\rho(x, y, \tau)} d\tau + \\ & + \int_0^t \frac{S_x(x, y, \tau)}{c\rho(x, y, \tau)} d\tau + \int_0^t \left[ \int_{\underline{x}(y)}^{\bar{x}(y)} c\rho(\xi, y, \tau) d\xi \right]^{-1} \times \\ & \times \left[ - \int_{\underline{x}(y)}^{\bar{x}(y)} S_x(\xi, y, \tau) d\xi - \bar{w}_x(\tau) + \underline{w}_x(\tau) \right] dt + \\ & + \int_0^t \frac{S_y(x, y, \tau)}{c\rho(x, y, \tau)} d\tau + \int_0^t \left[ \int_{\underline{y}(x)}^{\bar{y}(x)} c\rho(x, \eta, t) d\eta \right]^{-1} \times \\ & \times \left[ - \int_{\underline{y}(x)}^{\bar{y}(x)} S_y(x, \eta, \tau) d\eta - \bar{w}_y(\tau) + \underline{w}_y(\tau) \right] dt. \quad (5) \end{aligned}$$

Как было показано ранее, приведенная система интегральных уравнений (2)—(3) полностью эквивалентна дифференциальному уравнению теплопроводности (1).

Возможности организации нескольких параллельных вычислительных потоков связаны с возможностями решения системы интегральных уравнений (2)—(3). Заметим, что структура каждого из уравнений, входящих в систему, является однотипной, и поэтому для их решения естественно использовать единый алгоритм, обеспечивающий решение отдельного уравнения системы.

Для распараллеливания вычислений был предложен следующий алгоритм: если правые части уравнений (2) и (3) считать заданными известными функциями, то, как видно из структуры представленной системы, она распадается на систему независимых уравнений. Таким образом, каждое интегральное уравнение может быть решено независимо от остальных уравнений, что и позволяет организовать соответствующее число вычислительных потоков (в каждом из которых решается одно интегральное уравнение). После того как решения

каждого из уравнений найдены, вычисляются новые правые части системы и процедура повторяется. Блок-схема описанного здесь алгоритма представлена на рис. 1 (см. третью сторону обложки).

Этот алгоритм уже по первоначальному замыслу представляет собой итерационную процедуру, которая должна сходиться для того чтобы удовлетворялись согласованные условия по распространению теплоты по каждой из координатных осей. Поэтому представляется закономерным использовать эту же итерационную процедуру и для учета меняющихся с течением времени теплофизических характеристик. Таким образом, предлагаемый алгоритм представляет собой вполне естественную базу для решения нелинейных (по теплофизическим параметрам) задач теплопроводности. Отметим, что решение простых линейных задач теплопроводности хорошо выполняется известными методами, а использование многопроцессорных или многопоточковых алгоритмов оправдано именно при условии их нелинейности, когда другие вычислительные схемы не особенно эффективны. Это связано с тем, что решение отдельного интегрального уравнения, как правило, требует больших вычислительных затрат по сравнению с затратами при решении дифференциального уравнения.

Дополнительным преимуществом метода струн, кроме, как указано выше, возможности проводить параллельные вычисления, является именно их интегральная форма, позволяющая оценивать точность вычислений по текущей функции невязки. Это может быть актуальным при решении задач, в которых распределение температурных полей является критическим, другими словами, в задачах, когда меняющиеся со временем температурные режимы имеют первостепенное значение и влияют на эксплуатационные характеристики конечного изделия.

### Тестирование метода

Для эффективного решения исходной задачи, кроме возможности проведения параллельных вычислений, значение имеет и выбор метода решения отдельного интегрального уравнения. В литературе [например, 5, 6] имеется описание достаточно большого числа алгоритмов решения интегральных уравнений, например метод коллокаций или метод Галеркина [8], поэтому в данной работе остановимся именно на анализе предложенного алгоритма распараллеливания для метода струн.

Время решения всей задачи, как было отмечено выше, определяется временем решения наиболее трудоемкой из подзадач, выполняемых в отдельном вычислительном потоке. На рис. 2 (см. третью сторону обложки) показано сеточное разбиение и, соответственно, элементарные отрезки — "струны", на которых заданы интегральные уравнения. Можно заметить, что если область имеет сложную конфигурацию, то число узлов в "струне" меняется в за-

висимости от выбранного сечения. Очевидно, что наибольшее число узлов будет содержать наиболее пространственно протяженная "струна", и, следовательно, именно для нее решение интегрального уравнения будет наиболее затратным с вычислительной точки зрения.

Для тестирования описанного выше алгоритма была выбрана представленная на рис. 2 область, которая была аппроксимирована  $12 \times 10 = 120$  струнами. Очевидно, что минимальное вычислительное время при решении будут иметь задачи для струн, состоящих из трех точек, а максимальное время — для струны, состоящей из 10 точек. Для приведенного примера общее число интегральных уравнений составляет 120.

Для решения поставленной задачи написана программа на языке Visual C++, в которой на каждой итерации было организовано 120 вычислительных потоков (в одном потоке решалось одно интегральное уравнение). Теплофизические параметры были заданы в виде постоянных значений и с течением времени не менялись. Расчеты проводили на обычном ПК, при этом время решения составило менее 10 с. Это связано с тем, что при выбранном разбиении каждое интегральное уравнение приводило к системе, состоящей менее чем из 10 линейных алгебраических уравнений. Число итераций, обеспечивающих точность не менее 1 %, составляло 30—50, в зависимости от отрезка. Для сравнения скорости решения были проведены расчеты с использованием одного вычислительного потока, в котором последовательно решались те же 120 уравнений. Расчетное время в этом случае составляло 20 с.

Отсюда можно сделать вывод, что для задач небольшой размерности, когда время решения каждого отдельного интегрального уравнения не так велико, выигрыш распараллеливания получается незначительным — всего лишь в 2 раза. Это связано как с тем, что все вычисления выполняли на одном процессоре, так и с большими временами по выделению памяти и буферизации данных по сравнению с временем решения отдельной задачи. Поэтому значительного эффекта применения предложенного алгоритма можно ожидать лишь для трудоемких задач на больших сеточных разбиениях, когда число точек по каждой из осей составляет более 100, и при использовании многопроцессорных систем, например графических процессоров. Однако и здесь имеет смысл в каждом потоке решать не одно, а несколько интегральных уравнений.

### Заключение

Предлагаемый алгоритм параллельного решения нелинейных задач теплопроводности, основанный на интегральном представлении процесса теплопередачи — методе "струн", может быть эффективен для трудоемких задач большой размер-

ности. При этом может быть достигнута очень высокая степень параллелизма, и желательно, чтобы каждая из подзадач, входящих в вычислительный поток, содержала несколько "струн" для обеспечения сбалансированной загрузки процессоров.

#### Список литературы

1. Миллер Р., Боксер Л. Последовательные и параллельные алгоритмы: пер. с англ. М.: БИНОМ. Лаборатория знаний, 2006. 408 с.
2. Воеводин В. В., Воеводин Вл. В. Параллельные вычисления. СПб.: БХВ-Петербург, 2002, 608 с.
3. Эндриус Г. Р. Основы многопоточного, параллельного и распределенного программирования: пер. с англ. М.: Вильямс, 2003, 523 с.
4. Антонов А. С. Введение в параллельные вычисления (методическое пособие). М.: Изд-во МГУ, 2002. 69 с.

5. Барский А. Б. Параллельные процессы в вычислительных системах: планирование и организация. М.: Радио и связь, 1990. 255 с.
6. Барский А. Б. Параллельные технологии решения оптимизационных задач / Информационные технологии. Приложение. 2001. № 2. 32 с.
7. Демидович Б. П., Марон И. А., Шувалова Э. З. Численные методы анализа: приближение функций, дифференциальные и интегральные уравнения: учеб. пособие. / 3-е изд., перераб. М.: Наука, 1967. 368 с.
8. Власова Е. А., Зарубин В. С., Кувыркин Г. Н. Приближенные методы математической физики: учеб. для вузов / Под ред. В. С. Зарубина, А. П. Крищенко. М.: Изд-во МГТУ им. Н. Э. Баумана, 2001. 700 с.
9. Светушков Н. Н. Метод струн в задачах многомерной нестационарной теплопроводности // Информационные технологии. 2014. № 12. С. 14–19.
10. Светушков Н. Н. Возможности компьютерного моделирования технологических процессов термической обработки // Труды МАИ. 2012. № 58 ([www.mai.ru/science/trudy/](http://www.mai.ru/science/trudy/)).

N. N. Svetushkov, Associated Professor, Moscow Aviation Institute (National Research University), Moscow, Russia Associate Professor of department "Applied mathematics and physics", e-mail: svt.n.n@mail.ru

## Parallel Strings Method for the Numerical Solution of Nonlinear Heat Conduction Problem

Currently, more attention is paid to the use of multiprocessor systems in order to reduce the computational cost for modeling various processes for scientific and applied research. If we consider the partial differential equation, the use of simple explicit schemes allows one to effectively carry out the parallelization of the computational process, but for large periods of time there is the question of the accuracy of the results, which is associated with the convergence of direct circuits. If the problem is non-linear, for example, when the characteristics of the environment depend on the unknown quantities, there is a need to organize iterative procedures, and use a variety of implicit schemes for partial differential equations. In this case, the algorithms used are known to be poorly amenable to parallelization. All these problems are acute and when you try to simulate the processes of distribution of heat in a spatially extended and geometrically complex heterogeneous objects, thermal characteristics depend on the temperature (this class includes, for example, the classical Stefan problem). The author proposes to solve the problem of a significant reduction in the problem of computing time for solving such problems by providing a balanced multi-threaded computational procedure based on its proposed integrated "method of string." The article describes in detail the method of parallel computing threads for solution of nonlinear transient heat conduction problems on the basis of this method. The main advantage of the proposed algorithm is that the partial differential equations are reduced to an equivalent system of integral equations. There is a diagram of computing threads and it is shown that the greatest number of processes can be organized in the case where each thread is solved one integral equation. Integral description underlying parallel approach allows not only to more effectively simulate the processes of heat transfer by suppressing the oscillatory behavior of solutions in the case of large temperature gradients, but also a natural way to take into account changes over time thermal characteristics of the product. There is a test calculations for the model problem and it is shown that a significant effectiveness of the algorithm can be achieved for large-scale problems and in the case of multi-technology (including graphics cards). The developed algorithm can be used to calculate the temperature fields in the micro-TVEL, with thermal machining of heat-resistant steels, as well as the formation of products of the prepregs for aircraft and missile industry.

**Keywords:** numerical methods, parallel computing, software, simulation, the heat equation, integral equations, the accuracy of calculations

#### References

1. Miller R., Boxer L. *Algorithms sequential & parallel: Unified approach*. Cengage Learning, 2012, 408 p.
2. Voevodin V. V., Voevodin V. V. *Parallel'ny'e vy'chisleniya*. SPb.: BXV-Peterburg, 2002, 608 p.
3. Andrews R. *Foundations of Multithreaded, Parallel, and Distributed Programming*, University of Arizona, Instock, 2000, 664 p.
4. Antonov A. S. *Vvedenie v parallel'ny'e vy'chisleniya*. Moscow: Izd. MGU, 2002. 69 p.
5. Barskij A. B. *Parallel'ny'e processy' v vy'chislitel'ny'x sistemax: planirovanie i organizaciya*, Moscow: Radio i svyaz', 1990. 255 p.
6. Barskij A. B. *Parallel'ny'e tehnologii resheniya optimizacionny'x zadach. Informacionny'e tehnologii*, Prilozhenie, 2001, no. 2, 32 p.

7. Demidovich B. P., Maron I. A., Shuvalova E. Z. *Chislenny'e metody' analiza: priblizhenie funkcyj, differenciarny'e i integral'ny'e uravneniya*. Uchebnoe posobie. Moscow: Nauka, 1967, 368 p.
8. Vlasova E. A., Zarubin V. S., Kuvyrkin G. N. *Priblizhenny'e metody' matematicheskoj fiziki*. Ucheb. dlya vuzov / Pod red. V. S. Zarusubina, A. P. Krishhenko. M.: Izd-vo MGTU im. N. E'. Bauman, 2001, 700 p.
9. Svetushkov N. N. *Metod strun v zadachax mnogomernoj nestacionarnoj teploprovodnosti, Informacionny'e tehnologii*, 2014, no. 12, pp. 14–19.
10. Svetushkov N. N. *Vozmozhnosti komp'yuternogo modelirovaniya tehnologicheskix processov termicheskoj obrabotki, Trudy' MAI*. 2012, no. 58 ([www.mai.ru/science/trudy/](http://www.mai.ru/science/trudy/)).

# СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ CAD-SYSTEMS

УДК 004.272.2

**Р. А. Соловьев**, канд. техн. наук, нач. отдела, e-mail: ZF-Turbo@yandex.ru,  
**Д. В. Тельпухов**, канд. техн. наук, науч. сотр., **В. С. Рухлов**, аспирант, мл. науч. сотр.,  
**П. С. Поперечный**, аспирант  
Институт проблем проектирования в микроэлектронике РАН

## Особенности проектирования модулярных многовходовых сумматоров с помощью современных САПР

*Исследуются различные методы и подходы к проектированию многовходовых сумматоров по модулю, которые используются в задачах цифровой обработки сигналов и требуют максимальной производительности. Показано, что использование предлагаемых подходов позволяет достичь уменьшения критического пути на 33 % либо уменьшения занимаемой площади до 65 % в сравнении с использованием позиционных сумматоров. На основе сравнения характеристик различных реализаций мультиоперандных сумматоров, синтезированных в базе ПЛИС и заказных СБИС, был выявлен наиболее эффективный метод построения многовходовых модулярных сумматоров.*

**Ключевые слова:** система остаточных классов, многовходовый сумматор, ПЛИС, СБИС

### Введение

Зачастую при реализации устройств цифровой обработки сигналов, таких, например, как фильтр с конечной импульсной характеристикой (КИХ-фильтр) [1], появляется потребность в использовании многовходовых сумматоров. При проектировании модулярных вычислительных структур, соответственно, требуются модулярные многовходовые сумматоры (ММС). Двухвходовые модулярные сумматоры на данный момент подробно изучены [2]. Если для позиционных устройств современные системы автоматизированного проектирования (САПР) сами строят наиболее эффективные реализации, то для модулярных устройств требуется оценить, какая из множеств возможных реализаций является наиболее эффективной на практике. Этот факт является определяющим при разработке устройств, для которых требуется максимально высокая производительность.

В данной работе были исследованы четыре наиболее перспективных метода построения ММС. Для оценки их эффективности в двух технологических базисах (САПР Quartus II для ПЛИС Altera и Synopsys Design Compiler для заказных СБИС) были синтезированы сумматоры с различным числом входов. Приведены результаты синтеза и проведено сравнение с позиционными многовходовыми сумматорами большой разрядности.

Данная статья имеет следующую структуру. Сначала приводится информация о двухвходовых модулярных сумматорах и принципах их реализации. Далее предлагаются четыре метода для реали-

зации ММС. Приводятся детали экспериментов на ПЛИС и СБИС, а также графики и таблицы с результатами синтеза для разработанных схем. И наконец, в заключении делается вывод, какой именно тип ММС наиболее эффективен на практике с точки зрения быстроедействия и аппаратных затрат.

### 1. Модулярные сумматоры двух переменных

Модулярные сумматоры от двух переменных имеют простейшую структуру и наиболее эффективно строятся на базе обычных позиционных сумматоров с коррекцией выхода, если значение суммы превзошло или равно значению модуля. Типовой пример на языке проектирования аппаратуры Verilog можно записать так:

```
module sum2_mod_7 (in1, in2, out);
    input [2:0] in1;
    input [2:0] in2;
    output reg [2:0] out;
    wire [3:0] w;
    assign w = in1 + in2;
    always @ (*)
    begin
        if(w < 7)
            out <= w;
        else
            out <= w - 7;
        end
    endmodule
```

Существуют также полностью комбинационные реализации [3] и оптимизированные реализации сумматоров для специальных модулей вида  $2^n - 1$ ,  $2^n$ ,  $2^n + 1$  [4]. В данной статье это не является ключевым вопросом и можно выбрать любую реализацию двухвходового модулярного сумматора, которая наиболее эффективна для данной технологии. Многовходовые сумматоры, которые описаны ниже, строятся на базе двухвходовых.

### Многовходовые модулярные сумматоры

Теперь подробнее рассмотрим методы построения модулярных сумматоров большого числа слагаемых с помощью двухвходовых модулярных сумматоров, описанных выше.

**Модулярный пирамидальный сумматор.** Модулярный пирамидальный сумматор является ступенчатой пирамидальной структурой из двухвходовых сумматоров по основанию выбранного заранее простого числа. Для  $n$ -входового ММС первый уровень будет состоять из  $n/2$  сумматоров. На следующем уровне число двухвходовых модульных сумматоров вдвое меньше, чем на первом, так как его входами являются выходы сумматоров из первого уровня и т. д., пока не останется один выход. Пирамидальный сумматор на восемь входов по модулю  $p$  может быть описан следующей формулой:  $S = \|(A + B)_p + (C + D)_p\|_p + \|(E + F)_p + (G + H)_p\|_p$ . Здесь и в дальнейшем мы будем пользоваться свойством операции модуля, которое позволяет вносить знак модуля в арифметические выражения [5]:

$$|x \cdot y|_p = |x|_p \cdot |y|_p.$$

Схема пирамидального сумматора на восемь входов по модулю  $p$  представлена на рис. 1.

**Модулярный сумматор с коррекцией на выходе.** Модулярный сумматор с коррекцией на выходе построен на базе позиционного многовходового

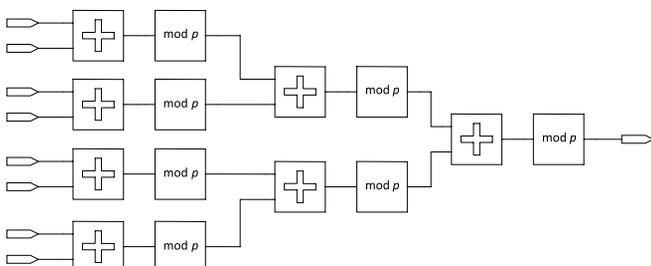


Рис. 1. Модулярный пирамидальный сумматор на восемь входов

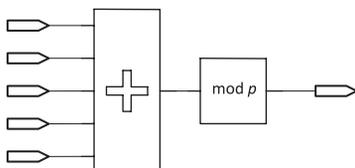


Рис. 2. Модулярный сумматор на пять входов с коррекцией на выходе

сумматора (рис. 2). Все слагаемые суммируются в позиционном виде, а приведение к необходимому основанию проводится на последнем этапе вычислений, благодаря эффективным методам нахождения остатка целочисленного деления [6]. Таким образом, максимально эффективно используются все возможности САПР по оптимизации позиционных структур. Математически такой вид модулярного сумматора можно описать следующей формулой:

$$S = |A + B + C + D + E|_p.$$

**Модулярный последовательный сумматор.** Модулярный последовательный сумматор представляет собой традиционную последовательную структуру из двухвходовых модулярных сумматоров, расположенных друг за другом (рис. 3). Здесь данные с двух входов суммируются на модулярном сумматоре, после чего к результату прибавляются данные со следующего входа с последующим преобразованием к модулярному виду и т. д. Последовательный модулярный четырехвходовый сумматор может быть описан следующей формулой:

$$S = \|(A + B)_p + C\|_p + D\|_p.$$

**Модулярный сумматор с промежуточной коррекцией.** Данный метод развивает идею метода с коррекцией на выходе и является его обобщением, позволяя приводить к заданному модулю не только в конце вычислений, но также и в промежуточных точках мультиоперандного сумматора. Крайними случаями данного метода являются пирамидальный модулярный сумматор, где модуль берется после каждого двоичного сумматора, а также сумматор с коррекцией на выходе, где только финальный результат позиционного суммирования корректируется по модулю (рис. 4).

Следует отметить, что для данного метода возможно большое число конфигураций с точки зре-

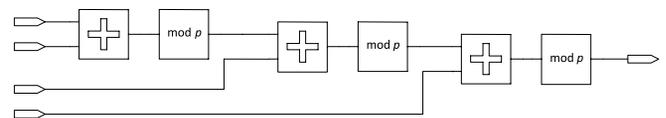


Рис. 3. Модулярный последовательный сумматор на четыре входа

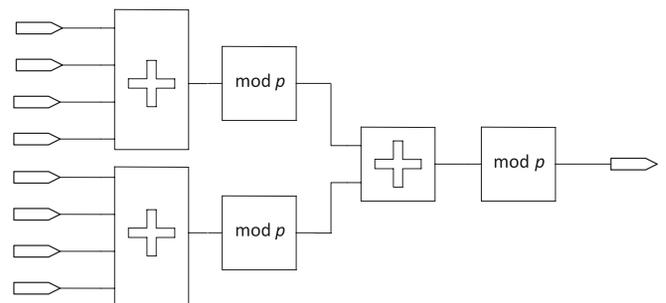


Рис. 4. Модулярный сумматор с промежуточной коррекцией (восемь входов, максимальное число элементов в позиционном сумматоре четыре)

ния количества и расположения промежуточных прямых преобразователей в вычислительном тракте мультиоперандного сумматора.

### Схема эксперимента

Для оценки временных и аппаратных затрат рассмотренных методов был реализован генератор функциональных Verilog описаний в целях определения наилучшего ММС, а также для сравнения с позиционными мультиоперандными сумматорами.

Для моделирования результата на ПЛИС использовали среду Quartus II Version 10,0 Build 218 06/27/2010 SJ Web Edition. В качестве используемой ПЛИС был выбран чип Altera Stratix II EP2S15F484C3. По результату компиляции оценивали число используемых вентилях, а также длину критического пути для каждого используемого метода.

Для проектирования в базе заказных СБИС использовали маршрут, включающий средства логического синтеза САПР Synopsys Design Compiler. Синтез проводили в базе библиотеки стандартных ячеек Nangate с проектными нормами 45 нм. Настройки синтезатора для всех рассматриваемых схем были выставлены на минимизацию длины критического пути [9–11].

Исследования проводили не только для различных типов сумматоров (модулярный пирамидальный сумматор, модулярный сумматор с коррекцией на выходе, модулярный последовательный сумматор, модулярный сумматор с промежуточной коррекцией и позиционный сумматор), но и для раз-

ного числа используемых входов (8, 16, 32, 64, 128). В качестве базового основания для всех модулярных схем был выбран семибитный модуль 113.

Для сравнения с модулярными сумматорами был выбран позиционный сумматор на 32 разряда, так как соответствующий динамический диапазон может быть покрыт основаниями, не превышающими 7 бит [7]. Позиционный многовходовый сумматор был описан на языке Verilog в наиболее общем виде, для того чтобы САПР могли синтезировать наилучшую архитектуру для заданных условий. Анализ результатов синтеза показал, что для реализации подобных схем современные САПР используют различные алгоритмы ускорения переносов.

### Результаты исследования

В табл. 1 представлены результаты тестов расчета задержки критического пути и число занимаемых сумматорами логических элементов на ПЛИС в зависимости от числа используемых входов.

Для наглядности здесь и далее полужирным выделены лучшие результаты для данного числа входов сумматора. На рис. 5 и 6 представлены графики зависимости длины критического пути и числа используемых элементов ПЛИС в зависимости от числа входов сумматора.

Аналогичные исследования были проведены для СБИС. Результаты можно увидеть на рис. 7, 8. В табл. 2 представлены сводные данные, полученные в результате моделирования.

Таблица 1

Сводная таблица полученных результатов на ПЛИС (критический путь (нс) и число элементов (шт.))

Тип сумматора	Число входов сумматора									
	8		16		32		64		128	
	нс	шт.	нс	шт.	нс	шт.	нс	шт.	нс	шт.
Пирамидальный	17,02	154	17,85	326	22,71	671	27,08	1392	29,57	2770
С коррекцией на выходе	16,57	<b>98</b>	18,41	<b>159</b>	23,38	<b>284</b>	<b>26,32</b>	<b>558</b>	<b>28,61</b>	<b>1023</b>
Последовательный	25,79	139	49,06	289	93,27	595	177,71	1235	364,91	2455
Позиционный 32 бит	<b>12,93</b>	147	<b>15,74</b>	365	<b>19,60</b>	729	29,84	1489	42,70	2945

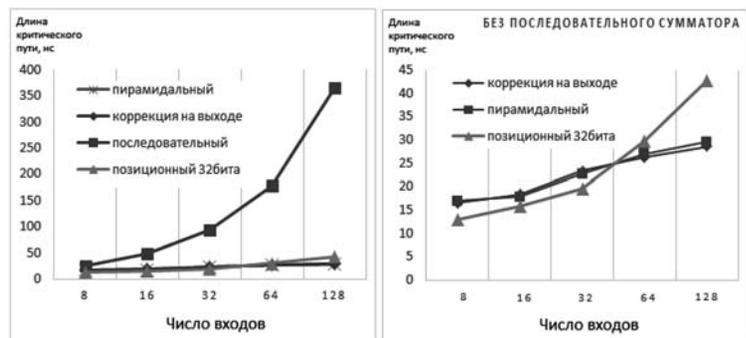


Рис. 5. График результатов моделирования критического пути в чипе ПЛИС в зависимости от используемого сумматора и числа его входов

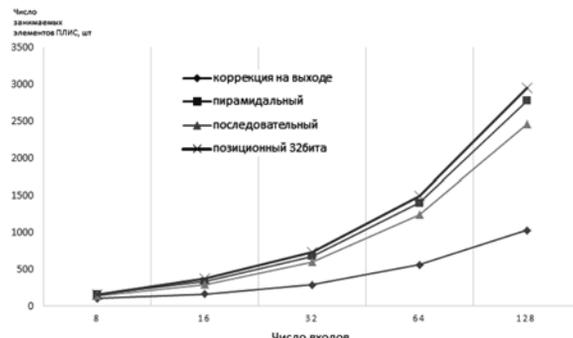


Рис. 6. Графики зависимости числа занятых блоков ПЛИС в зависимости от используемого сумматора и числа его входов

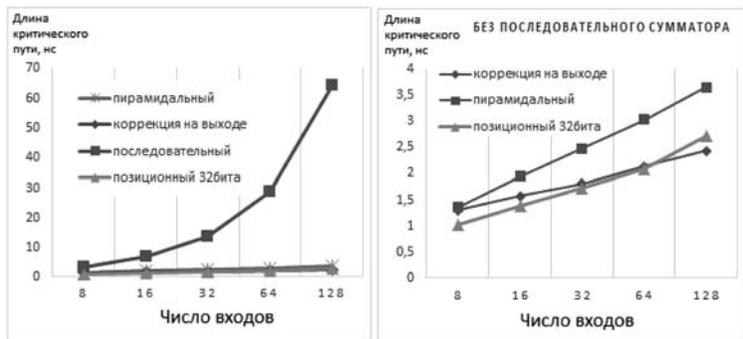


Рис. 7. График результатов моделирования критического пути в чипе СБИС в зависимости от используемого сумматора и числа его входов

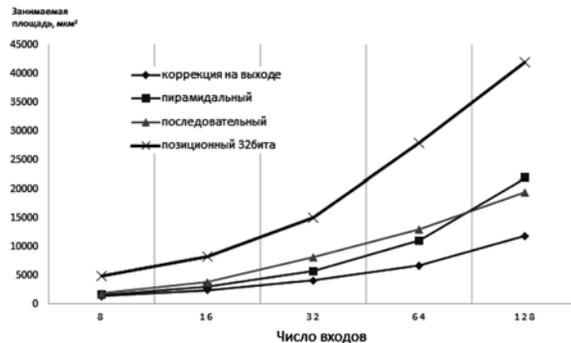


Рис. 8. Графики зависимости площади СБИС в зависимости от используемого сумматора и числа его входов

Сводная таблица полученных результатов для СБИС (критический путь (нс) и занимаемая площадь (мкм<sup>2</sup>))

Тип сумматора	Число входов сумматора									
	8		16		32		64		128	
	нс	мкм <sup>2</sup>	нс	мкм <sup>2</sup>	нс	мкм <sup>2</sup>	нс	мкм <sup>2</sup>	нс	мкм <sup>2</sup>
Пирамидальный	1,36	1671	1,94	3025	2,48	5751	3,03	11 102	3,65	21 906
С коррекцией на выходе	1,3	<b>1413</b>	1,57	<b>2420</b>	1,81	<b>4173</b>	2,14	<b>6737</b>	<b>2,43</b>	<b>11 860</b>
Последовательный	3,12	1871	6,71	3803	13,7	8139	28,78	12 966	64,29	19 444
Позиционный 32 бита	<b>1,03</b>	4974	<b>1,38</b>	8203	<b>1,73</b>	15 038	<b>2,09</b>	27 962	2,73	42 009

Таблица 2

Результаты исследования критического пути и занимаемых ячеек на ПЛИС при использовании промежуточной коррекции для различного числа входов (критический путь (нс) и занимаемая площадь (мкм<sup>2</sup>))

Число входов сумматора	8		16		32		64		128	
	нс	шт.	нс	шт.	нс	шт.	нс	шт.	нс	шт.
0										
2	<b>16,13</b>	239	20,15	510	26,44	1053	28,25	2169	34,58	4331
4	17,60	179	19,87	379	24,05	777	<b>25,78</b>	1584	28,62	3164
8	16,57	<b>98</b>	20,88	236	24,06	449	26,99	906	33,05	1820
16			<b>18,41</b>	<b>159</b>	24,09	358	28,59	725	30,19	1394
32					<b>23,38</b>	<b>284</b>	29,77	640	30,29	1225
64							26,32	<b>558</b>	30,30	1123
128									<b>28,61</b>	<b>1023</b>

Таблица 3

Результаты исследования критического пути и занимаемой площади на заказных СБИС при использовании промежуточной коррекции для различного числа входов (критический путь (нс) и занимаемая площадь (мкм<sup>2</sup>))

Число входов сумматора	8		16		32		64		128	
	нс	мкм <sup>2</sup>								
0										
2	1,31	1614	1,73	3642	2,45	6008	3,18	11893	3,62	24 038
4	<b>1,13</b>	1454	<b>1,39</b>	2810	1,95	4997	2,29	8210	2,82	16 082
8	1,3	<b>1413</b>	1,67	2587	2,2	4218	2,55	7596	3,2	13 928
16			1,57	<b>2419</b>	2,08	<b>3625</b>	2,58	6370	2,98	12 901
32					<b>1,81</b>	4172	2,28	<b>6225</b>	2,77	11 066
64							<b>2,14</b>	6736	2,63	<b>10 611</b>
128									<b>2,43</b>	11 860

Таблица 4

Получив результаты критического пути и площади при использовании каждого сумматора и проанализировав полученные результаты, было предложено дополнить исследование более подробным изучением сумматоров с коррекцией на выходе. Для этого на языке Verilog были реализованы различные варианты сумматоров с промежуточной коррекцией. Результаты проведенных исследований на ПЛИС представлены в табл. 3.

Аналогичные исследования были проведены для заказных СБИС. Результаты для СБИС приведены в табл. 4.

По результатам проведенных исследований для модулярных сумматоров с промежуточной коррекцией при реализации на ПЛИС и СБИС наблюдается тенденция улучшения параметров критического пути и площади при максимальном числе элементов в позиционном сумматоре, что соответствует выбранному ранее методу с коррекцией на выходе. Однако стоит учитывать, что выбирая оптимальный сумматор для конкретной задачи, в частных случаях можно получить улучшенные характеристики сумматора, используя метод промежуточной коррекции.

### Заключение

Налагаемые требования на разрабатываемое устройство диктуют выбор того или иного метода реализации мультиоперандного сумматора. В данной статье были рассмотрены различные методы реализации модулярных мультиоперандных сумматоров, а также проведены исследования их характеристик быстродействия и аппаратных затрат, по результатам которых можно сделать несколько выводов и рекомендаций:

- лучшим методом построения ММС как в отношении быстродействия, так и занимаемой площади является метод с коррекцией на выходе;
- большую часть всего модульного канала фактически занимает двоичный мультиоперандный сумматор, реализация которого эффективно отработана в современных САПР. Это объясняет наилучшие показатели в сравнении с другими методами реализации ММС;
- в некоторых частных случаях удается достичь некоторых улучшений при использовании метода промежуточной коррекции;
- использование модулярных мультиоперандных сумматоров, как и других модулярных вычислительных структур [1, 8], обычно целесообразно при реализации операций с большим вычислительным объемом. В приведенных эксперимен-

тах преимущество модулярных структур начинало проявляться в сумматорах с числом входов не менее 64.

Результаты демонстрируют схожие тенденции, как для СБИС, так и для ПЛИС, что подтверждает универсальность представленных выводов.

### Список литературы

1. **Тельпухов Д. В., Соловьев Р. А., Амербаев В. М., Балака Е. С.** Разработка аппаратного модулярного фильтра с конечной импульсной характеристикой на базе теоретико-числового быстрого преобразования Фурье // Всероссийская научно-техническая конференция "Проблемы разработки перспективных микро- и нанoeлектронных систем (МЭС)". Сб. трудов. 2014. № 4. С. 169—172.
2. **Амербаев В. М., Соловьев Р. А., Тельпухов Д. В.** Реализация библиотеки модульных арифметических операций на основе алгоритмов минимизации логических функций // Изв. Южного федерального университета. Технические науки. 2013. № 7 (144). С. 221—225.
3. **Dugdale M.** VLSI implementation of residue adders based on binary adders // IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing. 1992.
4. **Zimmerman R.** Efficient VLSI implementation of modulo  $2n \pm 1$  addition and multiplication // Proceedings 14<sup>th</sup> Symposium on Computer Arithmetic. 1999. P. 158—167.
5. **Omondi A., Premkumar B.** Residue Number Systems: Theory and Implementation. London: Imperial College Press. 2007. 312 p.
6. **Соловьев Р. А., Тельпухов Д. В.** Аппаратная реализация операции нахождения остатка целочисленного деления для входных данных большой разрядности в модулярной арифметике // Известия высших учебных заведений. Электроника. 2013. № 4 (102). С. 75—83.
7. **Амербаев В. М., Константинов А. В., Тельпухов Д. В.** Выбор технологичных модулей в задаче оптимизации модулярных умножителей и сумматоров // 16-я Всероссийская межвузовская научно-техническая конференция студентов и аспирантов "Микроэлектроника и информатика—2009". Сб. тезисов. М.: МИЭТ. 2009.
8. **Амербаев В. М., Соловьев Р. А., Тельпухов Д. В., Щелков А. Н.** Исследование эффективности модулярных вычислительных структур при проектировании аппаратных одноктактных умножителей // Известия Южного федерального университета. Технические науки. 2014. № 7 (156). С. 248—254.
9. **Стемпковский А. Л., Гаврилов С. В., Каграманян Э. Р.** Методы логико-временного анализа заказных блоков СБИС // Известия вузов. Электроника. 2008. № 5. С. 41—50.
10. **Гаврилов С. В., Гудкова О. Н., Егоров Ю. Б.** Методы ускоренной характеристики библиотек элементов СБИС с контролем заданной точности // Известия вузов. Электроника. 2010. № 3. С. 51—59.
11. **Гаврилов С. В., Гудкова О. Н., Егоров Ю. Б.** Статистический анализ сложных функциональных блоков // Известия вузов. Электроника. 2010. № 5. С. 58—64.

**R. A. Solovyev**, Ph. D., Head of the Department, e-mail: ZF-Turbo@yandex.ru,  
**D. V. Telpukhov**, Ph. D., Researcher of the Department, **V. S. Rukhlov**, Junior Researcher of the department,  
**P. S. Poperechny**, Postgraduate Student Institute for Design Problems in Microelectronics

## Study of Multi-Input Modular Adders Implemented on FPGA and VLSI by Modern CAD Tools

*In this paper different methods of the design of multi-input modular adders are studied, which are used in digital signal processing tasks and require maximal performance. Depending on the desired goal, up to 33 % reduction of the critical path and up to 65 % reduction in area can be achieved compared to designs using adders based on positional number systems. The most efficient method to implement multi-input modular adders was determined. As opposed to other papers, the research was carried out for the purpose of implementing adders on FPGA and custom VLSI for different numbers of inputs.*

**Keywords:** residue number system, multi-input adder, FPGA, VLSI

### References

1. **Telpuhov D. V., Solovyev R. A., Amerbaev V. M., Balaka E. S.** Razrabotka apparatnogo moduljarnogo fil'tra s konechnoj impul'snoj karakteristikoj na baze teoretiko-chislovogo bystrogo preobrazovanija Fur'e. *Vserossijskaja nauchno-tehnicheskaja konferencija "Problemy razrabotki perspektivnyh mikro- i nanojelektronnyh sistem (MJeS)".* Sb. trudov. 2014, no. 4, pp. 169–172.
2. **Amerbaev V. M., Solovyev R. A., Telpuhov D. V.** Realizacija biblioteki modul'nyh arifmeticheskijh operacij na osnove algoritmov minimizacii logicheskijh funkcij. *Izv. Juzhnogo federal'nogo universiteta. Tehniceskie nauki.* 2013, no. 7 (144), pp. 221–225.
3. **Dugdale M.** VLSI implementation of residue adders based on binary adders. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing.* 1992.
4. **Zimmerman R.** Efficient VLSI implementation of modulo  $2n \pm 1$  addition and multiplication. *Proceedings, 14<sup>th</sup> Symposium on Computer Arithmetic.* 1999, pp. 138–167.
5. **Omondi A., Premkumar B.** Residue Number Systems: Theory and Implementation. London: Imperial College Press. 2007. 312 p.
6. **Solovyev R. A., Telpuhov D. V.** Apparatnaja realizacija operacii nahozhdenija ostatka celochislennogo delenija dlja vhodnyh dan-

nyh bol'shoj razrjadnosti v moduljarnoj arifmetike. *Izv. vysshijh uchebnyh zavedenij. Jelektronika.* 2013, no. 4 (102), pp. 75–83.

7. **Amerbaev V. M., Konstantinov A. V., Telpuhov D. V.** Vybortehnologichnyh modulej v zadache optimizacii modul'nyh umnozhitel'ej i summatorov. *16-ja Vserossijskaja mezhvuzovskaja nauchno-tehnicheskaja konferencija studentov i aspirantov "Mikroelektronika i informatika-2009".* Sb. tezisov. M.: MIJeT. 2009.

8. **Amerbaev V. M., Solovyev R. A., Telpuhov D. V., Shhelokov A. N.** Issledovanie jeffektivnosti moduljarnykh vychislitel'nykh struktur pri proektirovanii apparatnykh odnotaknykh umnozhitel'ej. *Izv. Juzhnogo federal'nogo universiteta. Tehniceskie nauki.* 2014, no. 7 (156), pp. 248–254.

9. **Stempkovskij A. L., Gavrilov S. V., Kagramanian E. R.** Metody logiko-vremennogo analiza zakaznykh blokov SBIS. *Izv. vuzov. Elektronika.* 2008, no. 5, pp. 41–50.

10. **Gavrilov S. V., Gudkova O. N., Egorov Ju. B.** Metody uskorennoj karakterizacii bibliotek jelementov SBIS s kontrol'em zadannoj tochnosti. *Izv. vuzov. Elektronika.* 2010, no. 3, pp. 51–59.

11. **Gavrilov S. V., Gudkova O. N., Egorov Ju. B.** Statisticheskij analiz slozhnykh funkcional'nykh blokov. *Izv. vuzov. Elektronika.* 2010, no. 5, pp. 58–64.

УДК 004.942

**М. М. Гурарий**<sup>1</sup>, канд. техн. наук, ст. науч. сотр., e-mail: gourary@ippm.ru,  
**С. Г. Русаков**<sup>1</sup>, д-р техн. наук, гл. науч. сотр., e-mail: rusakov@ippm.ru,  
**Р. Р. Алиев**<sup>2</sup>, д-р техн. наук, вед. науч. сотр., e-mail: rubaliev@gmail.com

<sup>1</sup> Институт проблем проектирования в микроэлектронике РАН,

<sup>2</sup> Институт теоретической и экспериментальной биофизики РАН

## Применение методов схемотехнического моделирования для анализа биоэлектрических процессов в многоклеточных системах

*Показана возможность применения средств математического моделирования электронных схем для решения актуальной проблемы оценки воздействия внешних электрических полей на распределение потенциала внутри клетки и на поведение многоклеточных биологических систем. Обоснован выбор электрических эквивалентов внутриклеточных процессов. Обсуждаются результаты моделирования.*

**Ключевые слова:** клеточная ткань, мембрана, метод конечных объемов, передаточная функция, схемотехническое моделирование, уравнения Максвелла, электрическая схема

## Введение

Изучение электрической активности клеток и тканей является одной из важнейших проблем биологической физики. После создания Ходжкиным и Хаксли модели (модель Х—Х) распространения нервного импульса [1] большая часть теоретических исследований электровозбудимых клеток повторяла формализм Ходжкина—Хаксли. Расширения и модификации модели состояли в добавлении к ней дополнительных мембранных токов, обнаруженных в электровозбудимых клетках, таких как кардиомиоциты и нейроны [2]. Однако очевидным ограничением модели Х—Х при рассмотрении отдельных клеток, например кардиомиоцитов, является предположение об эквипотенциальности цитоплазмы клетки. Это предположение достаточно обосновано при рассмотрении низкочастотных процессов, присутствующих в норме при физиологических условиях. Для оценки частот стоит иметь в виду, что наиболее быстрый мембранный натриевый ток имеет характерное время около 1 мс (соответственно, речь идет о частотах около 1 кГц). Однако эквипотенциальность цитоплазмы клетки нарушается при наличии высокочастотных колебаний, например, при воздействии внешних высокочастотных электромагнитных полей либо при наличии быстроменяющихся полей, например, при разряде кардиостимулятора. В этом случае актуальным является изучение развития электропаразии ткани — процесса, обеспечивающего и сопровождающего успешную дефибрилляцию миокарда.

В данной работе рассмотрен подход, позволяющий моделировать распределение потенциала внутри клетки, в цитоплазме. Результаты моделирования обеспечивают связь с традиционными моделями типа модели Х—Х, описывающими мембранные токи.

В статье решение проблемы моделирования переноса заряда в клетках предлагается связать с адаптацией вычислительных процедур и средств схемотехнического моделирования. Мотивацией такого подхода является высокая степень автоматизации при моделировании электрических характеристик, достигнутая к настоящему времени, а также наличие универсальных алгоритмов, обеспечивающих широкий спектр видов моделирования во временной и в частотной областях.

Представление электрическими эквивалентами основных клеточных механизмов переноса было рассмотрено в работе [3]. Показаны вычислительные преимущества моделирования на уровне обыкновенных дифференциальных уравнений с применением законов Кирхгофа. Однако в работе [3] не представлены способы получения таких эквивалентных цепей. Настоящая статья направлена на установление связей исходных уравнений с возможными электрическими эквивалентами и дальнейшую разработку алгоритмов, обеспечивающих повышение производительности моделирования биоэлек-

трических процессов в многоклеточных структурах произвольной формы методами схемотехнического моделирования.

## Электрическая модель клеточной ткани

Электрически клеточную ткань можно представить как совокупность клеток, находящихся в однородной жидкой межклеточной среде внешнего электролита. Каждая клетка состоит из мембраны и содержащегося в ней внутреннего электролита — цитоплазмы. Толщина мембраны (6...8 нм) намного меньше размеров клетки (для кардиомиоцита 10...100 мкм).

Для анализа воздействия электрического поля на ткань была предложена методика моделирования на основе электрической модели клеточной ткани, названной "transport lattice" [3]. Эта методика получила дальнейшее развитие в работах [4—6]. В рамках модели клеточная ткань представляется в виде эквивалентной электрической схемы, состоящей из двухполюсников, соединяющих соседние узлы прямоугольной сетки.

Принцип построения эквивалентной схемы представлен на рис. 1 и может быть описан следующим образом. Если оба соединяемых узла находятся по одну сторону мембраны, т. е. принадлежат одному электролиту, то двухполюсник представляется RC-секцией, где сопротивление и емкость определяются соответствующими значениями для куба со стороной  $l$  (шаг сетки):

$$R_e = \rho_e / l, \quad C_e = \kappa_e \varepsilon_0 l, \quad (1)$$

где  $\rho_e$  — удельное сопротивление,  $\kappa_e$  — относительная диэлектрическая проницаемость соответствующего (1 — внешнего или 2 — внутреннего) электролита,  $\varepsilon_0$  — абсолютная диэлектрическая проницаемость вакуума.

Если соединяемые узлы находятся по разные стороны мембраны толщиной  $d$ , то двухполюсник между ними представляется в виде последовательного соединения трех секций:

а) две RC-секции с параметрами  $R_e/2$ ,  $2C_e$ , представляющие электрические характеристики полу-

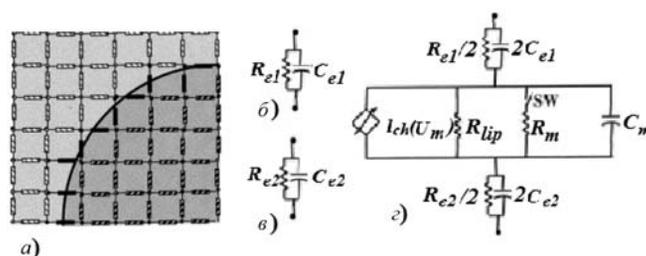


Рис. 1. Построение электрической схемы для моделирования воздействия поля на клеточную ткань:

а — общий вид схемы вблизи мембраны; б — двухполюсник между узлами внешнего электролита; в — двухполюсник между узлами внутреннего электролита; г — двухполюсник между узлами по разные стороны мембраны

кубов  $l \times l \times l/2$  внешнего и внутреннего электролитов;

б) секция, соответствующая мембране толщиной  $d$ , площадью  $l^2$  и напряжением  $U_m$ . Секция включает  $RC$ -цепь с параметрами

$$R_{lip} = \rho_{lip} d / l^2, \quad C_m = \kappa_m \varepsilon_0 l^2 / d, \quad (2)$$

а также подключенную параллельно с ней нелинейную модель ионных каналов с вольт-амперной характеристикой

$$I_{ch}(U_m) = U_m \left( \frac{g_m^+ + g_m^-}{2} \right) + w \left( \frac{g_m^+ - g_m^-}{2} \right) \operatorname{Incosh} \frac{U_m - U_{m0}}{w} + I_c, \quad (3)$$

и модель электропорации в виде последовательно-го соединения сопротивления и ключа, замыкание которого осуществляется при напряжении на мембране  $U_m = 0,5 \text{ V}$ :

$$R_m = 10^3 \rho_{e1} d / l^2 \text{ при } U_m \geq 0,5 \text{ V}. \quad (4)$$

В выражениях (2)–(4)  $\kappa_m, \rho_{lip}$  — относительная диэлектрическая проницаемость и постоянная составляющая удельного сопротивления мембраны;  $g_m^+, g_m^-$  — асимптотические (при прямом и обратном смещении) проводимости ионного канала;  $U_{m0}$  — потенциал покоя клетки;  $w = 22 \text{ мВ}$ ,  $I_c = 36 \text{ пА}$  — параметры аппроксимации.

Отметим, что представленные секции а) и б) соответствуют кубу со стороной  $l$ , который делится на две равные части мембраной толщиной  $d$ . Мембрана при этом расположена перпендикулярно направлению тока (рис. 2, а). Такое представление объясняется принятой в рамках модели ступенчатой аппроксимацией криволинейной поверхности клетки (рис. 2, б).

Отметим, что приведенный на рис. 1 вид модели в виде электрических эквивалентов носит эмпирический характер. Желательно, чтобы подобная модель была явно связана с исходными физическими процессами.

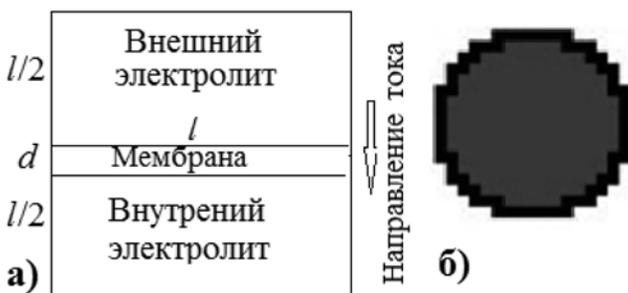


Рис. 2. Элементарная ячейка, соответствующая электрической ветви между узлами, находящимися по разные стороны мембраны (а); ступенчатая аппроксимация [3] криволинейной поверхности клетки (б)

## Обоснование метода

Для анализа биологических структур часто используется [7–9] модель на базе метода проводимостей (*Admittance Method*). Метод основан на разбиении ткани на параллелепипеды, которые представляются набором  $RC$ -цепочек, расположенных вдоль каждого из ребер (рис. 3).

Выражения для сопротивления и емкости параллелепипеда (с размерами  $\Delta X, \Delta Y, \Delta Z$ ) вдоль координатной оси имеют вид (например, для оси  $X$ ):

$$R_x = \rho_e \frac{\Delta X}{\Delta Y \cdot \Delta Z}, \quad C_x = \kappa_e \varepsilon_0 \frac{\Delta Y \cdot \Delta Z}{\Delta X}.$$

Эти выражения для куба со стороной  $l$  совпадают с формулой (1) для моделей на рис. 1, б, в, что свидетельствует о соответствии выражения (1) методу проводимостей. Однако участок ткани, включающий разные стороны мембраны (см. рис. 1, з), сгенерированный в соответствии с методом проводимостей, должен содержать  $RC$ -цепочки, параллельные мембране. Поскольку в модели на рис. 1 таких цепочек нет, такую модель нельзя считать полученной с помощью метода проводимостей.

Покажем далее, что для вывода модели клетки в виде электрических эквивалентов можно применить исходную модель на основе уравнений Максвелла для электрического поля в однородной среде:

$$\nabla j = 0; \quad (5a)$$

$$j = \sigma E + \varepsilon \frac{\partial E}{\partial t}; \quad (5b)$$

$$E = \nabla V. \quad (5в)$$

Здесь уравнение (5а) определяет закон сохранения заряда через нулевую дивергенцию плотности полного тока ( $j$ ), которая задается в (5б) как сумма

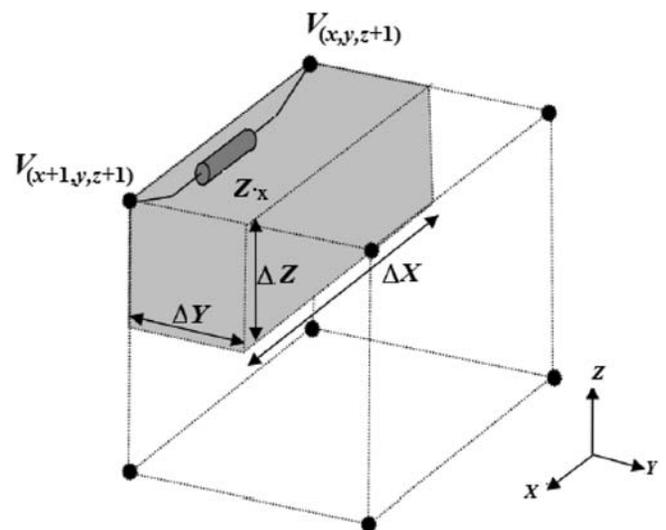


Рис. 3. Принцип формирования эквивалентной схемы для представления параллелепипеда в однородном участке ткани на основе метода проводимостей

тока проводимости и тока смещения, а уравнение (5в) определяет напряженность поля ( $E$ ) как градиент потенциала ( $V$ ).

Для численного решения системы (5) можно воспользоваться методом конечных объемов [10], который преобразует формулы (5) в систему алгебраических уравнений за счет разбиения пространственной области на множество элементарных объемов (ячеек) и выполнения следующих операций:

- интегрирование (5а) по объему ячейки и представление результата в соответствии с теоремой Остроградского—Гаусса в виде поверхностного интеграла от плотности тока;
- представление тока, протекающего через грань ячейки, как произведение плотности тока в центре грани на ее площадь;
- оценка плотности тока в центре грани на основе определения напряженности поля в этой точке заменой градиента в (5в) конечно-разностной формулой, использующей значения потенциалов в центрах ячеек.

Применение этих принципов для вычисления тока  $I_S$ , протекающего через грань ячейки площадью  $S$  (рис. 4), приводит к выражениям

$$E_S = \frac{V_2 - V_1}{l},$$

$$I_S = S j_S = \frac{S\sigma}{l} (V_2 - V_1) + \frac{S\varepsilon}{l} \frac{d(V_2 - V_1)}{dt}, \quad (6)$$

где  $E_S$  — напряженность поля в центре грани;  $V_1$ ,  $V_2$  — потенциалы соседних узлов;  $l$  — расстояние между узлами;  $j_S$  — плотность тока, протекающего через грань.

Из формулы (6) можно получить выражение для тока  $I_S$  в виде

$$I_S = G_{12}(V_2 - V_1) + C_{12} \frac{d(V_2 - V_1)}{dt}, \quad (7)$$

где

$$G_{12} = 1/R_{12} = S\sigma/l, \quad C_{12} = S\varepsilon/l. \quad (8)$$

Для  $RC$ -ячейки (7), представляющей куб со стороной  $l$ , имеем  $S = l^2$ , что приводит к совпадению формулы (8) с (1). Это соответствует параметрам двухполюсников между узлами модели внутри электролита (см. рис. 1, б, в). Для анализа куба со сто-

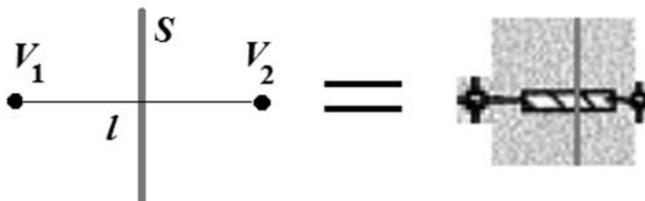


Рис. 4. Эквивалентная схема для представления тока, протекающего через грань ячейки

роной  $l$ , разделенного мембраной (см. рис. 2, а), необходимо его каждую половину рассмотреть как две ячейки. Одна ячейка с размерами  $l/2 \times l \times l$  занимает весь объем полукуба, а вторая прилегает к мембране и имеет бесконечно малую толщину с размерами  $0 \times l \times l$ . Тогда расстояние между узлами в центрах этих ячеек равно  $l/2$ , что соответствует параметрам  $RC$ -цепочек на рис. 1, з в направлении нормали к плоскости мембраны. В направлении, параллельном мембране,  $RC$ -цепочки отсутствуют в силу бесконечно малой толщины ячейки, приводящей к нулевым значениям проводимости и емкости.

Таким образом, мы показали, что модель в виде эквивалентной электрической цепи (см. рис. 1) может быть получена путем применения метода конечных объемов к уравнениям Максвелла в среде электролита.

### Модификация метода

Основным ограничением обсуждаемого метода является ступенчатая аппроксимация криволинейной поверхности клетки (см. рис. 2, б). Это приводит к неправильной оценке площади поверхности. Например, в двумерном случае отличие длины окружности диаметром  $D$  ( $L_{cir} = \pi D$ ) от длины ее ступенчатой аппроксимации, равной периметру описанного квадрата —  $L_{step} = 4D$ , определяется соотношением

$$L_{step}/L_{cir} = 4/\pi \approx 1,27. \quad (9)$$

В трехмерном случае отличие площади поверхности сферы диаметром  $D$  ( $L_{sph} = \pi D^2$ ) от площади ее ступенчатой аппроксимации, равной поверхности описанного куба  $L_{step} = 6D^2$ , определяется соотношением

$$L_{step}/L_{cir} = 6/\pi \approx 1,91. \quad (10)$$

Отметим, что соотношения (9), (10) не зависят от пространственного шага и поэтому не могут быть снижены за счет его уменьшения.

На рис. 5 показан другой пример существенной ошибки из-за ступенчатой аппроксимации пло-

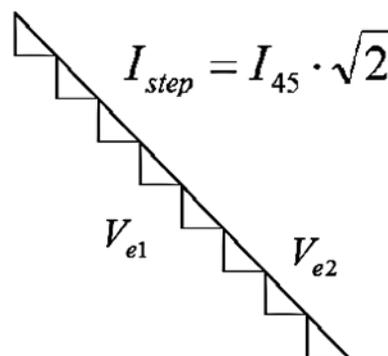


Рис. 5. Ошибка в расчете тока при ступенчатой аппроксимации плоской мембраны, проходящей под углом  $45^\circ$  к оси координат

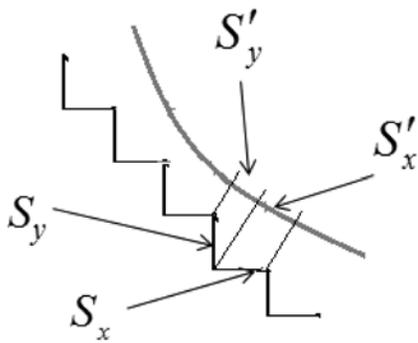


Рис. 6. Иллюстрация принципа замены фактической площади ступеньки на площадь ее проекции на исходную мембрану

ской мембраны, проходящей под углом  $45^\circ$  к оси координат. Оценка тока через мембрану при постоянных потенциалах с каждой стороны на основе ее ступенчатой аппроксимации ( $I_{step}$ ) превосходит истинное значение ( $I_{45}$ ) в 1,41 раза. Отметим, что ошибка не может быть снижена за счет уменьшения шага.

Для того чтобы исключить ошибку, связанную со ступенчатой аппроксимацией мембраны, не используя при этом сложных непрямоугольных сеток, предлагается следующий подход, основанный на модификации метода ступенчатой аппроксимации (рис. 6).

При оценке тока, протекающего через плоскость ступеньки, следует учитывать не ее фактическую площадь, а площадь ее проекции на исходную мембрану.

Другими словами, ток мембраны, протекающий через плоскость ступеньки, рассчитывается умножением плотности тока  $j_m$  не на площадь ступеньки  $S_{step}$  по формуле

$$I_{step} = j_m(V_m)S_{step},$$

а на площадь ее проекции на мембрану  $S'_{step}$  по формуле

$$I_{step} = j_m(V_m)S'_{step}.$$

Здесь  $S'_{step} = S_{step}(n_{step}n)$ , где  $n_{step}$  — единичный вектор нормали к поверхности ступеньки,  $n$  — единичный вектор нормали к поверхности мембраны. В частности, в двумерном случае  $n = [\cos\alpha, \sin\alpha]$ , где  $\alpha$  — угол между нормальными  $n_{step}$  и  $n$ , поэтому для плоскости ступеньки вдоль оси  $x$   $S'_x = S_x \cos\alpha$ , а для плоскости ступеньки вдоль оси  $y$   $S'_y = S_y \sin\alpha$ .

Для проверки предложенного усовершенствования метода была разработана вычислительная процедура. Процедура выполняет три варианта расчета сферической клетки между параллельными электродами:

- расчет на основе исходной модели (см. рис. 1);
- расчет на основе предложенной модификации;

- расчет с помощью теоретической модели для сферической клетки в идеальном однородном поле (формулы приведены в работе [11]).

Использовалась линейная модель мембраны, включающая только  $RC$ -цепочку (2). В качестве результатов моделирования бралась частотная передаточная функция от исходной напряженности поля к напряженности поля в верхней точке сферической мембраны, рассчитанная программой схемотехнического анализа.

Результаты расчетов представлены на рис. 7. Графики погрешностей, рассчитанных относительно теоретической зависимости, для исходной модели и для модели на основе предложенной модификации показаны на рис. 8.

Таким образом, из результатов вычислительных экспериментов видно, что предложенный подход существенно повышает точность модели.

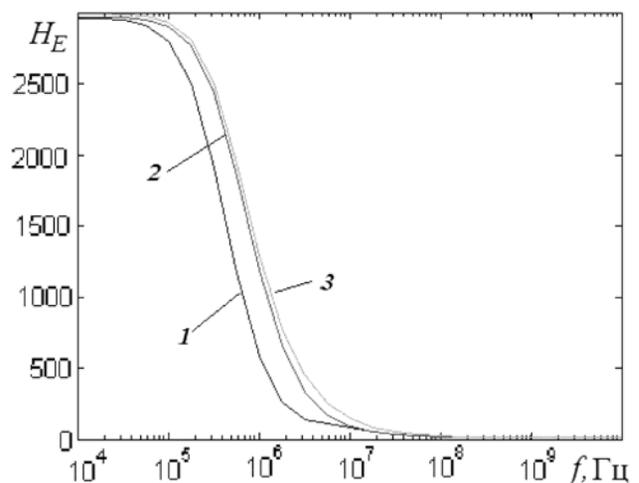


Рис. 7. Передаточная функция  $H_E$  по напряженности поля, рассчитанная на основе исходной модели (1), на основе предложенной модификации (2) и с помощью теоретической модели (3)

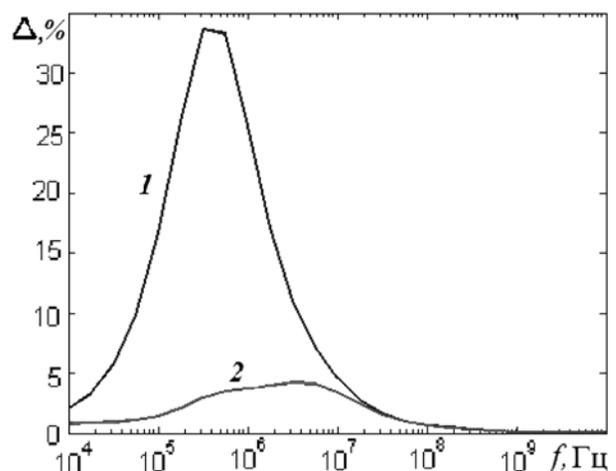


Рис. 8. Относительная погрешность (%) расчета передаточной функции на основе исходной модели (1) и на основе предложенной модификации (2)

## Численные эксперименты

Представленная модель клетки в форме электрических эквивалентов с предлагаемыми вычислительными модификациями была введена в программу схемотехнического анализа. На рис. 9 и 10 применение программы иллюстрируется примером моделирования поведения одиночной сферической клетки. Приведены рассчитанные характеристики в частотной области и во временной области.

На рис. 9 представлены результаты расчета частотных передаточных характеристик для линейной модели мембраны. В этом случае из модели (см. рис. 1) были удалены нелинейные элементы, описываемые формулами (3) и (4), моделирующие ионные каналы и процесс электропорации соответственно. На рис. 9, а показана рассчитанная характеристика для трехмерной модели сферической клетки с числом исходных пространственных точек для моделирования ( $29 \times 29 \times 29$ ). Для сравнения на рис. 9, б приведена вычисленная частотная характеристика при замене трехмерной модели на двумерную модель. Число исходных пространственных точек ( $29 \times 29$ ).

На рис. 10 показаны результаты применения программы схемотехнического анализа для моделирования во временной области полной нелинейной модели клетки в двумерном варианте. Пример иллюстрирует возможность расчета переходных процессов при импульсном воздействии. Из рис. 10 следует, что напряжение на мембране в высшей точке ограничено значением 0,5 В из-за резкого падения сопротивления мембраны при срабатывании ключа, отражающего электропорацию (4).

## Заключение

В работе рассмотрен метод анализа электрического поля в клеточной ткани на основе электрической сети из двухполюсных элементов. Приведено обоснование модели на основе применения метода конечных объемов к уравнениям Максвелла для электрического поля в однородной среде.

Показано, что основным ограничением модели является использование ступенчатой аппроксимации поверхности клетки, что приводит к погрешности, значение которой не может быть снижено за счет уменьшения шага сетки. Для устранения указанной погрешности предложен подход, основанный на проецировании плоскости ступеньки на поверхность мембраны и использовании площади полученной проекции вместо площади исходной плоскости в формуле для расчета тока, протекающего через мембрану. Проведенные вычислительные эксперименты показали, что использование предложенного подхода позволило уменьшить погрешность модели в 7...8 раз.

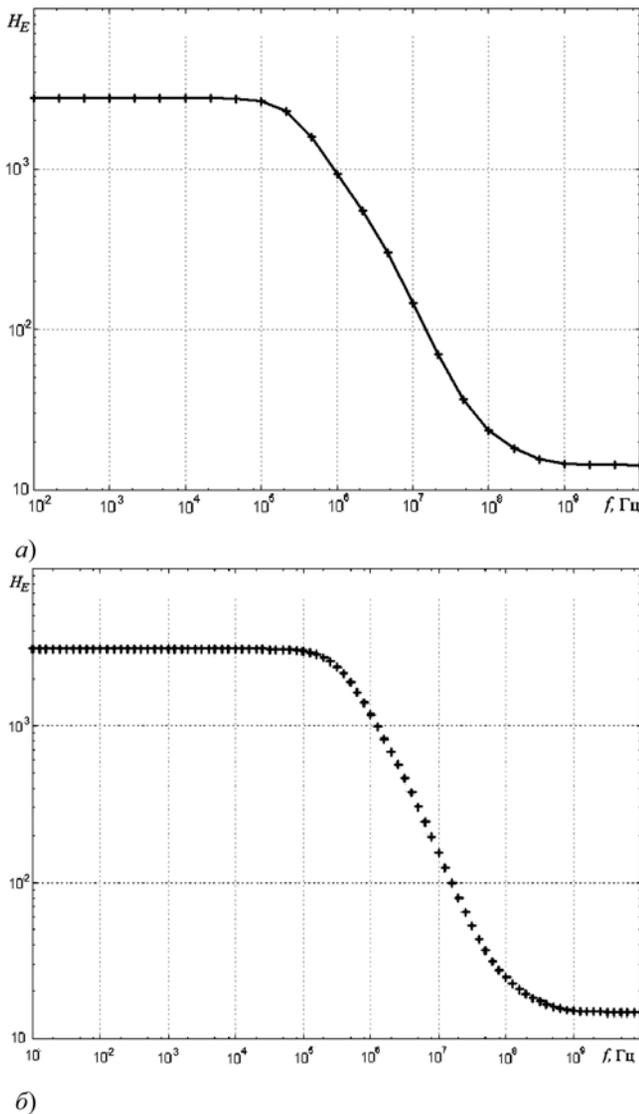


Рис. 9. Вычисленные частотные передаточные характеристики для трехмерной (а) и двумерной (б) моделей сферической клетки

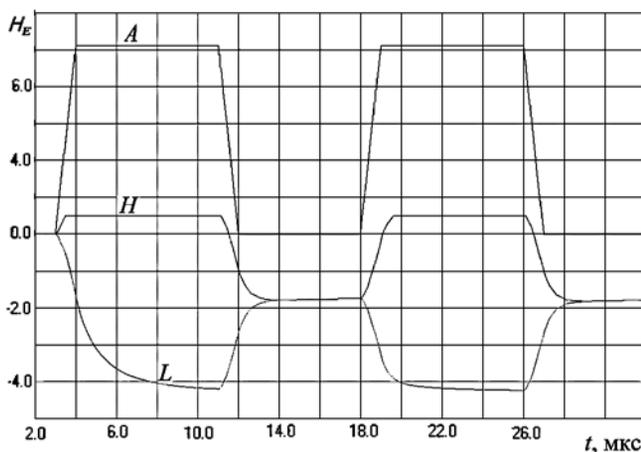


Рис. 10. Результаты моделирования во временной области воздействия импульсов на клеточную ткань. Входные (А) импульсы, графики относительной напряженности поля в верхней (Н) и нижней (L) точках мембраны

## Список литературы

1. **Hodgkin A. L. and Huxley A. F.** A quantitative description of membrane current and its application to conduction and excitation in nerve // *J. Physiol.* 1952. Vol. 117. P. 500–544.
2. **Алиев Р. Р.** Компьютерное моделирование электрической активности сердца // *Успехи физиологических наук.* 2010. Т. 41, № 3. С. 44–63.
3. **Gowrishankar T. R., Weaver J. C.** An approach to electrical modeling of single and multiple cells // *Proc. Natl. Acad. Sci. USA.* 2003. Vol. 100, N. 6. P. 3203–3208.
4. **Gowrishankar T. R., Stewart C., Weaver J. C.** Electroporation of a multicellular system: asymptotic model analysis // *Proc. of 26 Ann. Int. conf. of the IEEE Eng. Med. Biol. Soc.* 2004. P. 5444–5446.
5. **Stewart D. A., Gowrishankar T. R., Weaver J. C.** Cylindrical cell membranes in uniform applied electric fields: validation of a transport lattice method // *IEEE Trans. on Biomedical Engineering.* 2005. Vol. 52, N. 10. P. 1643–1653.
6. **Stewart D. A., Gowrishankar T. R., Weaver J. C.** Three dimensional transport lattice model for describing action potentials in

axons stimulated by external electrodes // *Bioelectrochemistry.* 2006. Vol. 69. P. 88–93.

7. **Armitage D. W., LeVeen H. H., Pethig R.** Radiofrequency-induced hyperthermia: Computer simulation of specific absorption rate distributions using realistic anatomical models // *Phys. Med. Biol.* 1983. N. 28. P. 31–42.

8. **Gandhi O. P., DeFord J. F., Kanai H.** Impedance method for calculation of power deposition patterns in magnetically induced hyperthermia // *IEEE Trans. Biomed. Eng. BME,* 1984. Vol. 31, N. 10. P. 644–651.

9. **Eberdt M., Brown P. K., Lazzi G.** Two-dimensional spicelinked multiresolution impedance method for low-frequency electromagnetic interactions // *IEEE Trans. Biomed. Eng.* 2003. Vol. 50, N. 7. P. 881–889.

10. **Рояк М. Э., Соловейчик Ю. Г., Шурина Э. П.** Сеточные методы решения краевых задач математической физики: учеб. пособие. Новосибирск: Изд-во НГТУ, 1998. 120 с.

11. **Kotnik T., Miklavcic D.** Second-Order Model of Membrane Electric Field Induced by Alternating External Electric Fields // *IEEE Trans. Biomed. Eng.* 2000. Vol. 47, N. 8. P. 1074–1081.

**M. M. Gourary<sup>1</sup>**, Senior Researcher, e-mail: gourary@ippm.ru,  
**S. G. Rusakov<sup>1</sup>**, Chief Researcher, e-mail: rusakov@ippm.ru,  
**R. R. Aliev<sup>2</sup>**, Leading Researcher, e-mail: rubaliev@gmail.com

<sup>1</sup> Institute for Design Problems in Microelectronics of Russian Academy of Sciences (IPPM RAS)

<sup>2</sup> Institute of theoretical and experimental biophysics of Russian Academy of Sciences

## Application of Circuit Simulation Methods for the Analysis of Bioelectrical Processes in Multicellular Systems

*The problem of the evaluation of the voltage distribution in biological tissues by circuit simulation tools is considered. The proposed approach is based on the transport lattice method. In this case a system of electrolytes, membranes, and electrodes is represented by an electrical network with the local charge transport or charge storage models.*

*In the paper the analysis of shortcomings of the transport lattice method is performed. The first shortcoming is the lack of the foundation of the method. It is shown that parameters of the network components can be obtained by the application of finite volume method to Maxwell equations. The boundary conditions are defined by membranes and electrodes. The main shortcoming of the transport lattice method is the error of the stepwise approximation of the membrane. It is shown that the value of the error cannot be decreased in this case by the grid stepsize reducing.*

*The approach to eliminate the error is proposed. It is based on the projection of the step plane onto the membrane surface and the application of the projection area instead of initial plane area in the expressions for the evaluation of membrane currents. The comparative simulations for traditional and proposed approaches are performed for the spherical cell in the homogeneous field. The errors are obtained by the comparisons with the known analytical expression. The simulation results confirm eightfold reducing of the error due to the proposed approach.*

**Keywords:** cellular tissue, cell membrane, finite volume method, transfer function, circuit simulation, electric network

### References

1. **Hodgkin A. L., Huxley A. F.** *J. Physiol.* 1952, vol. 117, pp. 500–544.
2. **Aliev R. R.** *Uspеhi fiziologicheskikh nauk*, 2010, vol. 41, no. 3, pp. 44–63 (in Russian).
3. **Gowrishankar T. R., Weaver J. C.** *Proc. Natl. Acad. Sci. USA.* 2003, vol. 100, no. 6, pp. 3203–3208.
4. **Gowrishankar T. R., Stewart C., Weaver J. C.** *Proc. of 26 Ann. Int. conf. of the IEEE Eng. Med Biol. Soc.* 2004, pp. 5444–5446.
5. **Stewart D. A., Gowrishankar T. R., Weaver J. C.** *IEEE Trans. on Biomedical Engineering.* 2005, vol. 52, no. 10, pp. 1643–1653.
6. **Stewart D. A., Gowrishankar T. R., Weaver J. C.** *Bioelectrochemistry*, 2006, vol. 69, pp. 88–93.

7. **Armitage D. W., LeVeen H. H., Pethig R.** *Phys. Med. Biol.*, 1983, no. 8, pp. 31–42.

8. **Gandhi O. P., DeFord J. F., Kanai H.** *IEEE Trans. Biomed. Eng., BME*, 1984, vol. 31, no. 10, pp. 644–651.

9. **Eberdt M., Brown P. K., Lazzi G.** *IEEE Trans. Biomed. Eng.* 2003, vol. 50, no. 7, pp. 881–889.

10. **Rojak M. Je., Solovejchik Ju. G., Shurina Je. P.** *Setochnye metody reshenija kraevykh zadach matematicheskoy fiziki: Uchebnoe posobie (Grid methods for boundary problems of mathematical physics: Tutorial)*, Novosibirsk: Izd-vo NGTU, 1998, 120 p. (in Russian).

11. **Kotnik T., Miklavcic D.** *IEEE Trans. Biomed. Eng.* 2000, vol. 47, no. 8, pp. 1074–1081.

**Л. Ф. Ноженкова**, д-р техн. наук, зав. отд., e-mail: expert@icm.krasn.ru,  
**О. С. Исаева**, канд. техн. наук, ст. науч. сотр., e-mail: isaeva@icm.krasn.ru,  
**Е. А. Грузенко**, аспирант, e-mail: Gruzenko@icm.krasn.ru,  
**Р. В. Вогоровский**, инженер, e-mail: vogorovskiy@icm.krasn.ru,  
**А. Ю. Колдырев**, инженер, e-mail: raventus@icm.krasn.ru,  
**А. А. Евсюков**, канд. техн. наук, ст. науч. сотр., e-mail: alev@icm.krasn.ru  
 Институт вычислительного моделирования СО РАН, Красноярск

## **Комплексная поддержка конструирования бортовых систем контроля и управления космических аппаратов на основе интеллектуальной имитационной модели**

*Предложена технология автоматизации и программной поддержки решения функциональных задач конструктора бортовой аппаратуры космического аппарата. Выполнен обзор состояния проблемы и существующих технологических и программных решений. Рассмотрены особенности информационного взаимодействия систем космического аппарата, осуществляющих контроль и командно-программное управление его основными системами. Выделены задачи моделирования бортовых систем и требования к интеграции, которым должна отвечать построенная модель. На основе системного подхода выполнено проектирование программного комплекса, предназначенного для информационно-графического и имитационного моделирования, анализа функционирования бортовой аппаратуры, а также для поддержки решения исследовательских задач. Апробация технологических подходов и программных решений выполнена на примере бортовой аппаратуры командно-измерительной системы. Результаты работы предназначены для внедрения в организацию, занимающуюся разработкой спутниковых систем.*

**Ключевые слова:** космический аппарат, бортовая аппаратура, командно-измерительная система, программно-математическая модель, имитационное моделирование, телекоманды, телеметрия, контрольно-проверочная аппаратура, поддержка проведения испытаний, анализ функционирования оборудования, учебно-исследовательская система

### **Введение**

Развитие высокотехнологичного и наукоемкого производства сложных программно-технических комплексов основывается на разработке методов интеллектуальной, информационной, графической поддержки конструирования и анализа функционирования технических устройств. Проведение экспериментальных исследований зачастую связано с экономическими и технологическими трудностями, что делает актуальной разработку единых подходов и программных решений, которые позволили бы представлять процессы функционирования и проводить анализ характеристик исследуемых объектов на программных моделях. Программная поддержка позволяет ускорить и значительно удешевить процессы конструирования различных технических устройств и систем.

Одним из научно-производственных направлений, для которого эффективно использование методов компьютерной поддержки и информационного моделирования, является конструирование и испытание бортовых систем контроля и управления космических аппаратов. В данном случае ограниченность возможностей экспериментального исследования космических аппаратов делает актуальной разработку методов программно-инструментальной

поддержки решения таких задач [1]. Существующее программное обеспечение, как правило, предназначено для оптимизации структуры, состава и размещения оборудования на борту космического аппарата, что позволяет найти решение, удовлетворяющее заданным требованиям по массогабаритным характеристикам, физическим свойствам и условиям эксплуатации [2]. Такого подхода оказывается недостаточно для моделирования объектов, имеющих собственную логику функционирования.

Особенностью рассматриваемых объектов анализа является наличие внутренней программы работы, собственных протоколов и наборов данных для информационного взаимодействия подсистем. Для автономных испытаний требуется выполнять моделирование передачи потоков данных от программ, имитирующих бортовые системы, к реальному оборудованию. Эти особенности должны учитываться при создании технологии моделирования и анализа функционирования. Программная поддержка конструирования таких объектов требует создания различных программных методов, позволяющих исследовать работу оборудования при различных заданных условиях. Решение исследовательских и конструкторских задач, возникающих на различных этапах проектирования, разработки и ввода в экс-

платацию бортовых систем, требует создания научно обоснованных методов, методик, рекомендаций и специального программного обеспечения.

### **Обзор существующих подходов к поддержке конструирования бортовой аппаратуры космического аппарата**

Стандарты Европейского космического агентства определяют системную разработку как междисциплинарный подход, управляющий всеми процессами преобразования технических требований в готовые решения в космической отрасли [3]. Стандартами определяются способы и форматы осуществления информационного взаимодействия бортовых систем между собой и с наземным комплексом управления. В процессе информационного взаимодействия применяются два типа данных: телекоманды (стандарт ESAPSS-04-107 [4]) и телеметрия (ESAPSS-04-106 [5]). Требуется разработать единую технологию моделирования и анализа функционирования бортовых систем космического аппарата, рассматривая их с точки зрения выполнения целевого назначения и абстрагируясь от внутреннего устройства исследуемого объекта, имитируя его поведение и информационное взаимодействие подсистем с участием указанных типов данных.

Основываясь на системном подходе, можно выделить основные задачи деятельности конструктора бортовой аппаратуры, требующие программного сопровождения и применения современных информационных технологий. К таким задачам относятся:

- моделирование особенностей функционирования бортовой аппаратуры в соответствии с различными условиями эксплуатации и назначением;
- проектирование основных принципов информационно-коммутиационного взаимодействия бортовых систем;
- проведение испытаний оборудования в составе контрольно-проверочной аппаратуры;
- анализ результатов испытаний;
- поддержка исследовательских функций и подготовка специалистов к работе с техническими устройствами и программным обеспечением.

Проведение различных имитационных экспериментов, изменение параметров функционирования модели, упрощение или детализация ее структуры будут способствовать лучшему пониманию свойств и характеристик бортовой аппаратуры [6].

Как правило, задачи поддержки конструирования рассматриваются в узконаправленных, не имеющих взаимосвязи исследованиях. Такие решения касаются отдельных проблем моделирования, испытания оборудования или анализа данных [7]. Для выбора метода и технологии решения поставленных задач рассмотрены существующие подходы, заключающиеся в применении специализированных программных систем [8], языков имитационного моделирования либо универсальных языков

программирования. Существует целый ряд программных систем как отечественного, так и зарубежного производства, предназначенных для выполнения структурного или имитационного моделирования. Например, можно отметить AnyLogic, Aimsun, Arena, Automod, Promodel [9]. Как правило, такие системы проблемно-ориентированные, содержат возможности сбора, обработки и вывода результатов моделирования, позволяют быстро и подробно анализировать предполагаемые исходы имитационных экспериментов [10]. Такие системы адаптированы под решение специфических задач, для которых они были разработаны, например, моделирование бизнес-процессов, логистика, организация транспортных потоков и пр. Однако их адаптация для моделирования функционирования бортовых систем накладывает существенные ограничения на постановку задачи, что делает невозможным построение адекватных моделей, отражающих ключевые аспекты контроля и управления космических аппаратов. Кроме того, затруднительно их применение для построения комплексных многофункциональных систем. Для обеспечения тесного взаимодействия между всеми программными модулями, поддерживающими различные задачи конструирования, необходимо создать средства имитационного моделирования как части единого программного обеспечения.

Для выбора методов программной поддержки анализа функционирования бортовой аппаратуры были рассмотрены существующие подходы к проведению испытаний технических систем. Под испытанием понимается экспериментальное определение количественных и (или) качественных свойств объекта путем воздействия на него контрольно-проверочной аппаратуры [11]. Строгое соответствие характеристик оборудования заданным на этапе проектирования техническим требованиям обеспечивает его правильную и долгосрочную работу на этапе эксплуатации, что имеет существенное значение в космической отрасли.

Испытания бортового оборудования должны проводиться при контроле телеметрической информации, отражающей его состояние. Достижение точности, достоверности и скорости получения результатов предъявляет дополнительные требования к программному обеспечению [12]. Анализ функционирования бортовой аппаратуры проводится как в составе борта космического аппарата, так и автономно. В этом случае требуется имитировать информационное взаимодействие бортовых систем и наземного комплекса управления для подтверждения заданных требований функционирования. Существующие специализированные аппаратно-программные комплексы обеспечивают высокие точность и скорость проведения измерений и позволяют управлять различным технологическим оборудованием [13].

Исследование подобного класса систем показало, что они направлены на решение отдельных задач испытания и анализа, не предназначены или имеют существенные ограничения по расширению измерительных функций. Такие системы реализованы в различных программных средах, что затрудняет, а в каких-то случаях делает невозможным интеграцию и совместное использование разработок различных производителей. Программное обеспечение должно основываться на современных технологиях разработки информационно-измерительных систем, соответствовать повышенным требованиям к измерительным возможностям оборудования, обеспечивать расширяемость и функциональную интеграцию с инструментами имитационного моделирования.

Таким образом, требуется разработать подход, позволяющий не просто объединить набор отдельных решений, но и разработать целостный интегрированный программный комплекс, предназначенный для моделирования, проектирования, проведения испытаний, анализа, обработки данных и сопровождения исследовательских задач.

### Системный анализ функциональных задач поддержки конструирования

Для решения задач комплексной поддержки конструирования применяли методы системного проектирования. Выполнен совместный анализ объекта исследования и процесса его создания и функционирования с учетом особенностей взаимодействия и взаимосвязей отдельных объектов и их частей как между собой, так и с внешней средой. В основе процесса поддержки конструирования должна лежать модель бортовой аппаратуры космического аппарата. Задача построения адекватной модели требует создания графической конфи-

гурации и информационного описания структурных элементов, коммутационных связей, параметров и характеристик. Функции модели определяются назначением исходного объекта моделирования и различаются в зависимости от задач бортовой аппаратуры и условий ее эксплуатации. Не существует однозначного описания алгоритмов работы бортовых систем. Требуется разработать такие способы задания поведения модели, которые позволили бы конструктору в простой и наглядной форме моделировать различные реализации ее функционирования.

Для решения этой задачи, как правило, применяют технологии искусственного интеллекта. Требуется разработать базу знаний, позволяющую задавать правила функционирования бортовой аппаратуры, создавать удобные инструменты ведения базы, выполнять разработку методов логического вывода. Это позволит изменять состояние модели, основываясь на знаниях и проектных решениях конструктора, имитируя функционирование реальной системы. Если для построения модели использовать технические описания готового оборудования, то она может лечь в основу процесса проведения его испытаний. Применение имитационного моделирования дает возможность детально изучать различные вопросы конструирования бортовых систем: опыт создания, способы информационного взаимодействия, особенности подготовки и проведения испытаний. Кроме того, моделирование позволит решать учебно-исследовательские задачи и, в свою очередь, будет способствовать повышению квалификации и качества подготовки персонала.

На основе концептуального описания задач конструирования разработаны функциональные диаграммы (рис. 1).

В связи со сложностью объекта исследования, отсутствием в открытом доступе подобных решений, которые бы могли лечь в основу системного проекта, процесс создания функциональных диаграмм получился итерационным. Применение итерационного принципа проектирования позволило последовательно приблизиться к выделению требований к программному обеспечению, детально описать необходимые данные, зависимые процессы, архитектуру и единые принципы создания программного обеспечения в целом, способы организации отдельных подсистем [14].

На основе детализации функциональных диаграмм выполнено проектирование подсистем программного комплекса в следующем составе:

- подсистема информационно-графического моделирования. Ее ос-



Рис. 1. Функциональная диаграмма поддержки конструирования

- новное назначение — формирование модели в соответствии с конфигурацией бортовой аппаратуры;
- интеллектуальная подсистема. Предназначена для задания логики работы бортовой аппаратуры в виде правил изменения состояний информационно-графической модели;
  - подсистема имитационного моделирования. Позволяет проводить имитационные эксперименты для анализа функционирования бортовой аппаратуры при различных условиях;
  - учебно-исследовательская подсистема. Предназначена для исследования работы бортовых систем в целях повышения квалификации;
  - подсистема проведения испытаний. Ее назначение — анализ характеристик работы оборудования в соответствии с техническими требованиями и сценариями функционирования.

Для оптимального распределения функций между программными системами описаны организационные и информационные ресурсы, выполнена унификация задач, разработаны модели данных и принципы организации единого информационного пространства, необходимые как для работы отдельных подсистем, так и для проекта в целом. Единое информационное пространство обеспечивается за счет создания базы данных, хранилища данных, базы знаний, содержащих описания параметров, режимов функционирования, правил поведения элементов модели, сценариев и результатов испытаний оборудования.

Результатом реализации системного проекта стал программный комплекс, предназначенный для поддержки конструирования бортовых систем управления, контроля параметров космического аппарата. Каждая подсистема предоставляет специализированный интерфейс, отражающий функции, соответствующие решаемым задачам.

### Информационно-графическое моделирование

Разработана среда информационно-графического моделирования, предназначенная для создания графического описания объекта моделирования, его составных частей и их взаимных связей. Степень детализации модели определяется целью моделирования. Для построения модели разработаны визуальные инструменты (рис. 2).

С помощью программного обеспечения конструктор имеет возможность создавать собственный проект аппаратуры или строить модель готового оборудования для анализа его функционирования, определять коммутационные интерфейсы, вводить

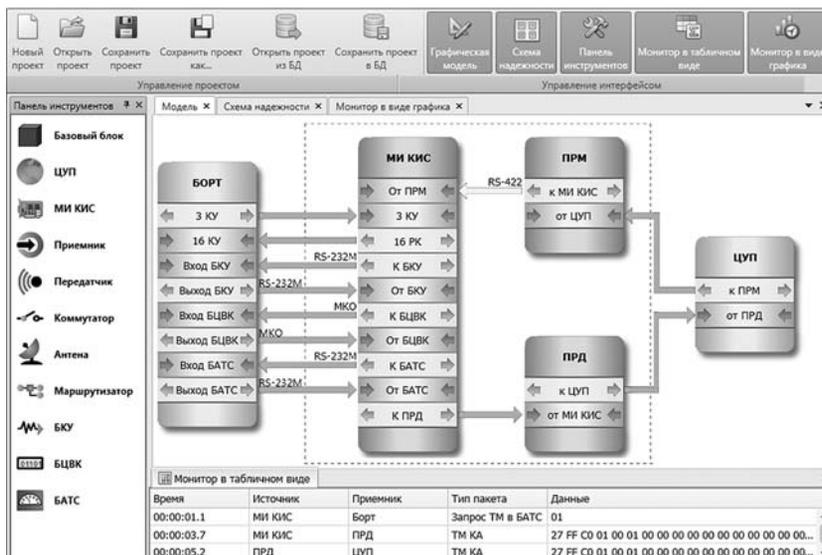


Рис. 2. Подсистема информационно-графического моделирования

параметры функционирования и задавать различные характеристики. Программное обеспечение позволяет манипулировать графическими объектами и характеристиками predetermined элементов из панели инструментов. Разработаны информационные модели бортовой аппаратуры, из которых конструктор может создавать собственные элементы, имеющие необходимую конфигурацию.

Задачи, возникающие перед конструктором в процессе моделирования, могут потребовать многократного изменения графической модели, упрощения или детализации отдельных функциональных блоков. Например, для анализа прохождения информации через определенный цифровой интерфейс модель упрощается — в нее могут быть включены связи только по выбранному интерфейсу. Информационно-графическая модель предназначена для проведения имитационных экспериментов.

### Интеллектуальные методы имитационного моделирования

Для создания методов работы имитационной модели выполнен переход от содержательного к формальному описанию процесса функционирования объекта моделирования на основе информационно-графической модели и базы знаний. База знаний содержит правила:

- информационного взаимодействия;
- работы отдельных элементов бортовой аппаратуры;
- выполнения командно-программного управления;
- прохождения пакетов телекоманд и телеметрической информации;
- описания реакций на изменения наблюдаемых параметров.

Правила позволяют манипулировать множествами входных воздействий, воздействий внешней среды, внутренних параметров и выходных наблюдаемых параметров.

Правила представляют собой конструкцию вида: "Если условие то действие". Левая часть правила задает условие его выполнения, а правая часть — действия, осуществляющие изменение состояния модели. Задание правил выполняется через визуальный редактор, предоставляющий простой и наглядный интерфейс для формирования символьных конструкций, описывающих логику работы модели (рис. 3).

В базе правил поддерживается модульность, которая позволяет вносить изменения знаний (добавляя, удаляя или модифицируя отдельные производные правила) без изменения логики работы независимых блоков. Выбранная структура знаний облегчает их интерпретацию и позволяет в естественном виде манипулировать графическими блоками модели и их информационными характеристиками.

База знаний применяется для интеллектуальной поддержки имитационного моделирования. Имитация функционирования модели осуществляется путем пошагового или автоматического изменения ее состояния в соответствии с заданными в базе знаний правилами. Для этого в процессе логического вывода выбираются правила, применимые к текущему состоянию модели, выполняются действия и визуализируется выполнение команд, имитируются процессы формирования и передачи пакетов данных. Имитационные эксперименты способствуют глубокому пониманию системы и позволяют представлять, как она реагирует на внешние воздействия, нештатные ситуации и как развивается с течением времени [15].

Программное обеспечение имитирует и визуализирует изменение состояния и активности элементов модели, формирование и передачу пакетов

данных через коммутационные интерфейсы. В графическом представлении каждая передача данных сопровождается анимационными эффектами. Интерактивный процесс моделирования позволяет конструктору проводить имитационные эксперименты, изменяя параметры приема, передачи сигналов, времени ожидания ответов и квитанций, переключая активность коммутационных интерфейсов и элементов модели для анализа ее поведения при нештатных ситуациях.

Если построение имитационной модели выполняется в соответствии с технической документацией на уже созданное оборудование, то такая модель является основой для проведения ее испытаний. В этом случае характеристики элементов модели предназначены для оценки параметров функционирования оборудования, контроля результатов испытаний на основе эталонных критериев, формирования реакций при выходе за граничные условия, а также формирования рекомендаций.

### Подготовка и проведение испытаний

Комплексная поддержка проведения испытаний бортовых систем космического аппарата основывается на исследовании технологических подходов к задаче построения измерительных информационных систем. Под измерительной системой понимается совокупность функционально объединенных средств измерений, вычислительной техники и вспомогательных устройств, взаимодействующих между собой и предназначенных для выработки и анализа измерительной информации о физических величинах, свойственных объекту [12]. Измерительные системы имеют характерные особенности, которые учитывались при проектировании и создании программного обеспечения: разнотипность измерительного оборудования и проводимых экспериментов, многократное изменение методики испытаний, расширяемость функций и набора исследуемых величин и др. Основные информационные процессы в такой системе — измерение и контроль, анализ и диагностика, сбор и хранение, отображение и обобщение. Все измерительные задачи выполняются с помощью контрольно-проверочной аппаратуры. Для взаимодействия с ней существуют специальные библиотеки программ. Эти программы составляют нижний уровень программного обеспечения, могут быть реализованы в разных средах разработки, работать по собственным принципам и быть заменяемыми в случае необходимости. Разрабатываемое программное обеспечение использует предоставляемые библиотеки и осуществляет поддержку измерительных задач, организуя верхний уровень взаимодействия пользователя и объекта контроля. Назначение подсистемы проведения испытаний заключается в управлении ходом испытаний в соответствии с определенными пользовательскими сценариями. Сценарии проведения испытаний

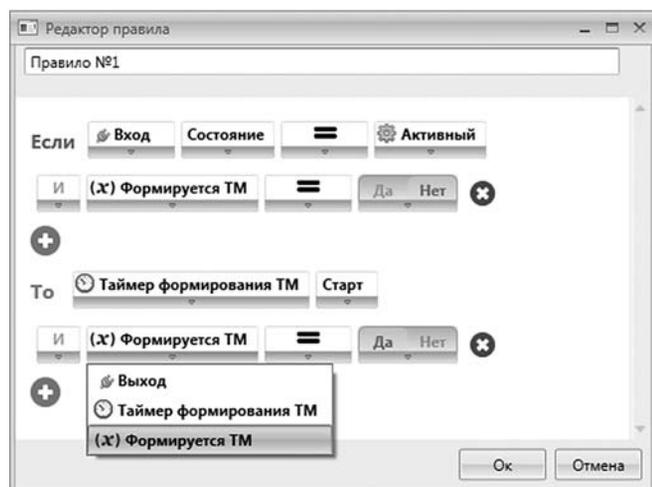


Рис. 3. Редактор правила

представляют собой набор измерений и действий на основе команд и параметров оборудования (рис. 4).

Процесс создания сценария — это строго упорядоченная последовательность шагов. На первом шаге формируется набор действий, которые представляют собой методы работы с измерительным оборудованием. Конструктор из действий формирует последовательности, представляющие собой задания на измерения и наблюдения, из которых строятся сценарии испытаний. Например, в качестве действий могут выступать: измерение номинальной частоты, определение частоты и частотной нестабильности несущего колебания, измерение спектра с использованием маски. Действия объединяются в задание. Построенные задания объединяются в сценарий, который позволяет анализировать состояние объекта испытаний. Для действий определяются условия функционирования, наблюдаемые и контролируемые параметры, а также способы визуализации результатов испытаний.

Программное обеспечение проводит проверку корректности испытаний в соответствии с техническими параметрами оборудования, заданными в информационно-графической модели. Функции анализа измерений формируют реакции при выходе за граничные условия, генерацию отчетов и рекомендаций по результатам проведенных испытаний. Поддерживается программная управляемость, т. е. возможность программным путем изменять состояние оборудования или алгоритмы его функционирования (в допустимых пределах), для оперативного изменения возможности всей системы в зависимости от конкретных задач или результатов анализа.

### Учебно-исследовательские функции

Учебно-исследовательская подсистема (рис. 5) представляет собой программный комплекс, предназначенный для интерактивного обучения и приобретения знаний, навыков работы инженерно-технического персонала и студентов в объеме, необходимом для проектирования и тестирования составных частей комплекса бортовой аппаратуры командно-измерительной системы.

В основе учебного курса лежит имитационная модель, позволяющая наглядно демонстрировать различные особенности функционирования бортовых систем космического аппарата. Модель может использоваться в каче-

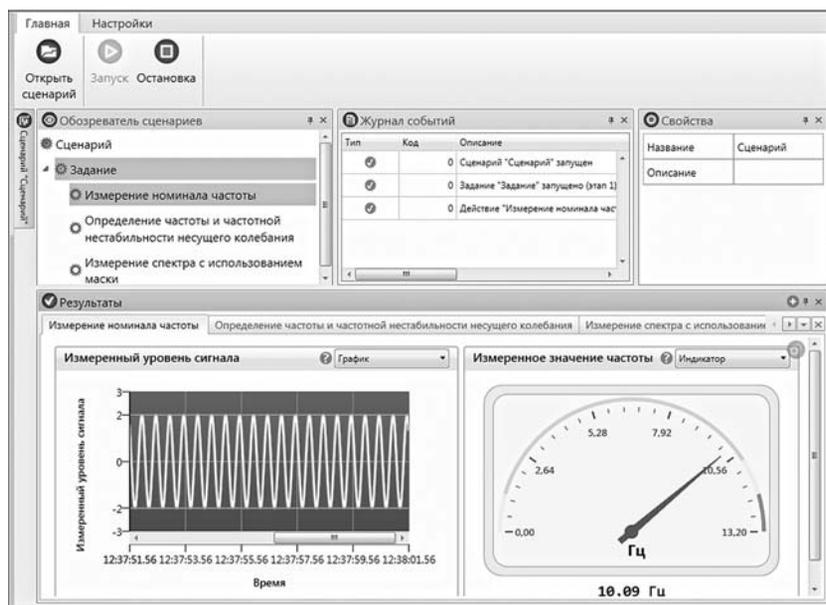


Рис. 4. Подсистема проведения испытаний и анализа результатов

стве тренажера для задания учебно-тренировочных сценариев. Поддерживаются задачи базовой подготовки персонала [16], выработки умений решения типовых задач, формирования навыков работы с оборудованием, а также действий в нестандартных ситуациях. Трансляция содержательного наполнения учебного курса осуществляется в соответствии с тематическими разделами либо согласно индивидуальным предпочтениям пользователей.

Функции моделирования обучающих ситуаций позволяют в графической нотации создавать модели процессов работы бортовой аппаратуры, управлять параметрами модели и вносить корректирующие поправки в обучающие сценарии. Возможность проводить модельные эксперименты способствует развитию способностей к проектированию, исследованию и анализу функционирования оборудования. Объединение информационных технологий транс-

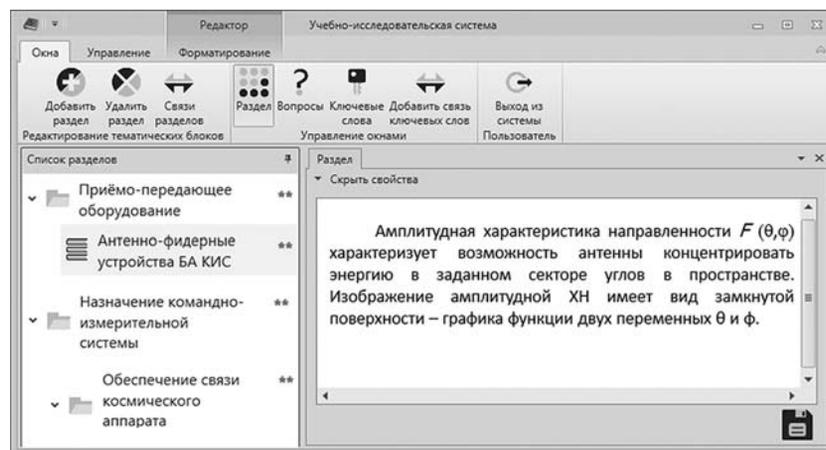


Рис. 5. Учебно-исследовательская подсистема

ляции учебного материала с функциями тренажера обеспечивает новые возможности подготовки специалистов.

Для организации индивидуальных траекторий обучения задаются информационные зависимости между учебными материалами, определяющие последовательности разделов, необходимые для успешного освоения каждой темы. Дополнительная навигация задается средствами ведения тематического тезауруса. Он объединяет в себе расширенный справочник ключевых слов, описывающих зависимые понятия и тематические разделы. Поддерживаются функции подготовки тематических заданий и выполнения контроля знаний. Обеспечивается возможность формирования подсказок и рекомендаций на основе информационных зависимостей характеристик обучающихся и индивидуальных траекторий обучения, что соответствует требованиям, предъявляемым к современным образовательным ресурсам.

### **Поддержка конструирования командно-измерительной системы космического аппарата**

Разработанные подходы и программное обеспечение были апробированы на примере решения задач комплексной информационно-графической поддержки конструирования бортовой аппаратуры командно-измерительной системы космического аппарата. Командно-измерительная система представляет собой одну из важнейших бортовых систем. Она предназначена для измерения параметров движения космического аппарата, приема и передачи различных видов информации, формирования и передачи на космический аппарат команд и программ управления, стандартных частот и сигналов времени для синхронизации работы бортового комплекса управления [17].

Построенная модель командно-измерительной системы содержит более 10 базовых элементов, включая резервные комплекты. Для задания информационного обмена в модели создано порядка 40 коммутационных соединений, имеющих различные типы интерфейсов: RS-232, RS-422, MKO, Spicewire. Основу модели составляют функциональные элементы специального назначения (приемник, передатчик, интерфейсный модуль командно-измерительной системы), описывающие возможности внутрисистемного информационного обмена. Внешнее взаимодействие командно-измерительной системы задается с помощью блоков бортовых систем космического аппарата: бортового комплекса управления, бортового цифрового вычислительного комплекса и бортовой аппаратуры телекоммуникации. Реализовано моделирование внешнего управления в части взаимодействия наземно-

го комплекса с командно-измерительной системой космического аппарата.

Инструменты графического моделирования позволяют выполнять детализацию или расширение модели в зависимости от решаемых задач и целей моделирования. На основе технических документов и спецификаций оборудования разработана база знаний, описывающая функционирование командно-измерительной системы космического аппарата. Информационное взаимодействие бортовых систем между собой и с наземным комплексом управления осуществляется на основе двух типов данных: телекоманд и телеметрической информации. Моделируются процессы получения телекоманд из наземного комплекса управления для имитации командно-программного управления бортовыми системами. Сведения о состоянии бортовых систем и квитанции о получении и выполнении команд управления содержатся в пакетах телеметрической информации, которые предназначены для передачи с борта космического корабля в наземный комплекс управления. В модели имитируется состав телеметрической информации, способы ее формирования и передачи на уровне информационного обмена.

Разработана база знаний, описывающая правила и порядок выбора комплекта рабочего оборудования, различных режимов, скорости передачи данных между бортовыми системами по выбранному коммутационному интерфейсу. Правила задают способы приема-передачи телеметрической информации, контроля состояния бортовой аппаратуры. Например, правило выбора скорости формирования телеметрической информации в основном комплекте интерфейсного модуля командно-измерительной системы имеет вид:

**Если** Команда = U10002N

**или** Команда = Рестарт

**То** Скорость формирования телеметрии =  
= 1000 бит/с

База знаний позволяет выполнять моделирование свойств и характеристик бортовой аппаратуры командно-измерительной системы в объеме, необходимом для проведения исследований структуры, состава, функций и области изменения параметров модели. Простой и наглядный способ описания правил позволяет конструктору самостоятельно представлять различные варианты работы оборудования.

Конфигурация модели полностью повторяет схему реального оборудования командно-измерительной системы и используется для проведения испытаний информационного взаимодействия между оборудованием командно-измерительной системы и бортовыми системами в случае, когда последние представлены своей имитационной моделью. Такое мо-

делирование позволяет проводить анализ функционирования аппаратуры с учетом изменения ограничений и целевых функций в условиях имитации взаимодействия с внешней средой. Проводимые во время испытаний проверки позволяют экспериментально подтвердить соответствия характеристик командно-измерительной системы установленным техническим требованиям.

Выполняется первичное наполнение информационно-справочного материала учебно-исследовательской подсистемы. Применение разработанных программных инструментов позволило реализовать различные обучающие ситуации, демонстрирующие логику работы бортовой аппаратуры командно-измерительной системы.

Результаты работы предназначены для внедрения в организации, занимающуюся разработкой спутниковых систем.

### Заключение

Результатом проведения системных исследований стала единая технология организации комплексной поддержки конструирования бортовых систем контроля и управления космического аппарата. В ее основе лежат общие принципы построения интеллектуальной информационно-графической модели, предназначенной для имитационного моделирования и анализа работы бортовой аппаратуры.

Предложенные методы и технологические подходы реализованы в программном обеспечении, которое позволяет выполнять моделирование архитектуры и логики функционирования в соответствии с назначением и условиями работы реального оборудования. Применение созданного программного обеспечения расширило возможности проведения экспериментальных исследований, предназначенных для выполнения анализа параметров оборудования и его соответствия техническим требованиям. Программное обеспечение выполнено на современном техническом уровне в соответствии с передовыми тенденциями развития информационных, измерительных и образовательных технологий.

В настоящий момент закончена разработка основных подсистем программного обеспечения. Апробация готовых программных решений выполнена на примере задачи поддержки конструирования бортовой аппаратуры командно-измерительной системы космического аппарата.

Готовится проведение испытаний программного обеспечения и внедрение его в опытную эксплуа-

тацию в организации-разработчике спутниковых систем. Применение программного комплекса даст возможность принимать научно обоснованные конструкторские решения на всех этапах жизненного цикла разработки бортовой аппаратуры.

*Работа выполнена при финансовой поддержке Минобрнауки России в Институте вычислительного моделирования СО РАН (договор № 02.G25.31.0041).*

### Список литературы

1. Александровская Л. Н., Круглов В. И., Кузнецов А. Г. Теоретические основы испытаний и экспериментальная отработка сложных технических систем. М.: Логос, 2003. 736 с.
2. Higdon K. P., Klaus D. M. Multidisciplinary design optimization for the development of human spacecraft // Conference Proceeding Paper "Earth & Space 2008: Engineering, Science, Construction, and Operations in Challenging Environments". March 3–5, 2008, Long Beach, California, United States. 2008. P. 1–10.
3. Space engineering. System engineering general requirements ECSS-E-ST-10C. Netherland: ESA publication division, 2009. 100 p.
4. Packet Telecommand Standard (ESA PSS-04-107). Is. 2: European space agency. 1992. 166 p.
5. Packet Telemetry Standard (ESA PSS-04-106). Is. 1: European space agency, 1988. 73 p.
6. System Design, Modeling, and Simulation using Ptolemy II: Ptolemy.org, 2014. URL: <http://ptolemy.org/systems>.
7. Фраленко В. П. Методы и алгоритмы обработки потоков данных в многопроцессорных вычислительных комплексах командно-измерительных систем: Автореферат дис. ... канд. техн. наук. Переславль-Залесский, 2011. 21 с.
8. Лычкина Н. Н. Технологические возможности современных систем моделирования // Банковские технологии. 2000. Вып. 9. С. 60–63.
9. Зайцева Н. О. Имитационное моделирование средствами системно-объектного подхода // Научные ведомости Белгородского государственного университета. 2012. Вып. 7–1. Т. 22. С. 155–159.
10. Советов Б. Я., Яковлев С. А. Моделирование систем. М.: Высшая школа, 2009. 343 с.
11. ГОСТ 16504–81. Испытания и контроль качества продукции. Основные термины и определения. М.: Стандартинформ, 2011. 24 с.
12. Раннев Г. Г. Измерительные информационные системы. М.: Издательский центр "Академия", 2010. 336 с.
13. Системы тестирования современной электроники. URL: [http://www.pegasuslogic.ru/docs/semiconductor\\_test.pdf](http://www.pegasuslogic.ru/docs/semiconductor_test.pdf).
14. Маклаков С. В. Создание информационных систем с All-Fusion Modeling Suite. 2-е изд., доп. М.: Диалог-МИФИ, 2007. 400 с.
15. Замятина О. М. Моделирование систем: учеб. пособие. Томск: Изд-во ТПУ, 2009. 204 с.
16. Башмаков А. И., Башмаков И. А. Разработка компьютерных учебников и обучающих систем. М.: Филинь, 2003. 616 с.
17. ГОСТ Р 53802–2010. Системы и комплексы космические. Термины и определения. М.: Стандартинформ, 2011. 28 с.

L. F. Nozhenkova, Head of Department, e-mail: expert@icm.krasn.ru,  
O. S. Isaeva, Senior Researcher, e-mail: isaeva@icm.krasn.ru,  
E. A. Gruzenko, Post-Graduate, e-mail: gruzenko@icm.krasn.ru,  
R. V. Vogorovskiy, Engineer, e-mail: vogorovskiy@icm.krasn.ru,  
A. Yu. Koldyrev, Engineer, e-mail: raventus@icm.krasn.ru,  
A. A. Evsyukov, Senior Researcher, alev@icm.krasn.ru  
Institute of computational modelling SB RAS, Russia, Krasnoyarsk

## Complex Support of Designing Onboard Control and Management Systems of Spacecrafts on the Basis of Intellectual Simulation Model

*This article describes the program-instrumental support of designing of the spacecraft onboard monitoring and control systems. The main purpose of this work is to automate all stages of the development life cycle of spacecraft onboard systems including designing, production, testing and commissioning. It proposed integrated approach of combining informational, measuring, computing and control functions within a single software. This software is a complex of four subsystems. The first is the graphic interfaces subsystem that allows to construct and to configure the model of the onboard monitoring and control systems. The subsystem of intelligent simulation modelling is intended to simulate process of passing of the information packets through the model of onboard monitoring and control system. The test execution subsystem provides complex support of the spacecraft systems testing by organizing test sequences and analyzing tests results. The educational system is intended for personnel training in the area of designing and testing of the spacecraft onboard monitoring and control systems. Proposed technological and software methods allow to make scientifically based design solutions at all stages of development life cycle of the spacecraft onboard systems. Proposed approach is applies to designing of the spacecraft command and measuring system that solves tasks of command control and monitoring of the spacecraft onboard systems.*

**Keywords:** spacecraft, onboard equipment, command and measuring system, program and mathematical model, simulation modeling, telecommands, telemetry packets, test and control equipment, test execution support, analysis of the functioning, educational and research system

### References

1. Aleksandrovskaja L. N., Kruglov V. I., Kuznecov A. G. *Teoreticheskie osnovy ispytanij i jeksperimental'naja otrabotka slozhnyh tehnikeskikh sistem*. M.: Logos, 2003. 736 p.
2. Higdon K. P., Klaus D. M. Multidisciplinary design optimization for the development of human spacecraft. *Conference Proceeding Paper "Earth & Space 2008: Engineering, Science, Construction, and Operations in Challenging Environments"*. March 3–5, 2008, Long Beach, California, United States. 2008, pp. 1–10.
3. *Space engineering. System engineering general requirements ECSS-E-ST-10C*. Netherland: ESA publication division, 2009. 100 p.
4. *Packet Telecommand Standard (ESA PSS-04-107)*. Iss. 2: European space agency. 1992. 166 p.
5. *Packet Telemetry Standard (ESA PSS-04-106)*. Iss. 1: European space agency. 1988. 73 p.
6. *System Design, Modeling, and Simulation using Ptolemy II*: Ptolemy.org, 2014. URL: <http://ptolemy.org/systems>.
7. Fralenko V. P. *Metody i algoritmy obrabotki potokov dannyh v mnogoprocessornyh vychislitel'nyh kompleksah komando-izmeritel'nyh sistem*: Avtoreferat kand. tehn. nauk. Pereslavl'-Zalesskij, 2011. 21 p.
8. Lychkina N. N. Tehnologicheskie vozmozhnosti sovremennyh sistem modelirovaniya. *Bankovskie tehnologii*. 2000, vol. 9. P. 60–63.
9. Zaitseva N. O. Imitatsionnoe modelirovanie sredstvami sistemno-ob'ektnogo podhoda. *Nauchnye vedomosti Belgorodskogo Gosudarstvennogo universiteta*. 2012, vol. 22, iss. 7–1, pp. 155–159.
10. Sovetov B. Ja., Jakovlev S. A. *Modelirovanie sistem*. M.: Vysshaja shkola, 2009. 343 p.
11. **GOST 16504–81**. *Ispytanija i kontrol' kachestva produkcii. Osnovnye terminy i opredelenija*. M.: Standartinform, 2001. 24 p.
12. Rannev G. G. *Izmeritel'nye informacionnye sistemy*. M.: Izdatel'skij centr "Akademija", 2010. 336 p.
13. *Sistemy testirovaniya sovremennoj elektroniki*. URL: [http://www.pegasuslogic.ru/docs/semiconductor\\_test.pdf](http://www.pegasuslogic.ru/docs/semiconductor_test.pdf).
14. Maklakov S. V. *Sozdanie informacionnyh sistem s AllFusion Modeling Suite*. 2-e izd., dop. M.: Izdatel'stvo Dialog-MIFI, 2007. 400 p.
15. Zamjatina O. M. *Modelirovanie sistem: uchebnoe posobie*. Tomsk: Izd-vo TPU, 2009. 204 p.
16. Bashmakov A. I., Bashmakov I. A. *Razrabotka komp'juternyh uchebnikov i obuchajushih sistem*. M.: Informacionno-izdatel'skij dom "Filin", 2003. 616 p.
17. **GOST R 53802–2010**. *Sistemy i komplekсы kosmicheskie. Terminy i opredelenija*. M.: Standartinform, 2011. 28 p.

УДК 004.912

А. Г. Яшина, ассистент, e-mail: ayashina.pmi@gmail.com,  
Д. Е. Прозоров, д-р техн. наук, профессор, e-mail: prozorov.de@gmail.com  
Вятский государственный университет, г. Киров

## Взвешенная косинусная мера векторной модели информационного поиска речевых документов

*Рассмотрена задача поиска речевых документов по текстовому запросу. Предложено использование взвешенной косинусной меры сходства в качестве функции релевантности векторной модели информационного поиска. Веса вычисляются посредством нечеткого сравнения слов на основе длины наибольшей общей подстроки, либо расстояния Левенштейна. Проведен анализ эффективности поиска речевых документов в зависимости от используемых алгоритмов сравнения слов.*

**Ключевые слова:** информационный поиск, речевые документы, векторная модель информационного поиска, косинусная мера

### Введение

Развитие информационных технологий привело к увеличению количества электронных документов различного типа. Существующие системы информационного поиска ориентированы в основном на обработку текстовых документов [1, 2]. Однако к настоящему времени широкое распространение получили мультимедийные документы, в которых существенная часть информации передается посредством речи. Примерами таких документов являются радио- и видеонews, аудиокниги, записи лекций, доклады конференций и т. п. Документы, содержащие только речь (без музыки), называются речевыми [3].

Поиск речевых документов по текстовому или устному запросу относится к области "Spoken Document Retrieval" (SDR), которая находится на стыке распознавания речи и информационного поиска. Основное отличие поиска речевых документов от текстового поиска заключается в наличии ошибок распознавания, которые искажают изначальное содержание речевых документов, что отражается на эффективности поиска. Поэтому актуальной задачей является разработка методов вычисления оценки релевантности речевых документов запросу, в которых учитываются ошибки распознавания.

Методы поиска речевых документов по текстовому запросу с учетом ошибок распознавания можно разделить на две группы [4]. К первой группе отно-

сятся методы, основанные на распознавании слитной речи в тексте и применении алгоритмов поиска по тексту [5, 6]. Методы второй группы используют фонемное распознавание или транскрибирование речевых документов [7–10]. Существуют комбинированные методы, объединяющие оба подхода [11, 12].

Выделяют [1] следующие модели информационного поиска: булевы, векторные, вероятностные. Модель информационного поиска описывает "представление" содержания документов и запросов пользователя, а также метод вычисления оценки релевантности документа запросу. Векторную модель широко применяют при информационном поиске по тексту. Основные достоинства данной модели заключаются в простоте и наличии различных методов взвешивания слов и вычисления оценки релевантности. Документы и запросы в векторной модели представляются векторами пространства, базисные векторы которого соответствуют словам, входящим в содержание документов. Оценка релевантности документа запросу вычисляется как значение меры близости двух векторов.

В качестве меры близости широко используют косинусную меру близости (косинус угла между векторами) [1] на основе точного сравнения слов.

Существуют также модифицированные меры близости на основе нечеткого сравнения строк. Например, в работе [13] вводится "мягкая" косинусная

мера (*soft cosine measure*), применяемая для классификации текстов:

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\sum_{i=0}^N \sum_{j=0}^N s_{i,j} a_i b_j}{\sum_{i=0}^N \sum_{j=0}^N s_{i,j} a_i a_j \cdot \sum_{i=0}^N \sum_{j=0}^N s_{i,j} b_i b_j},$$

где  $s_{i,j}$  — мера сходства слов сравниваемых текстов. "Мягкая" косинусная мера близости позволяет сравнивать содержание текстов в векторном пространстве, базис которого не является ортогональным.

Метод, описанный в работе [13], при сравнении текстов учитывает сходство слов, которое может быть "по смыслу", полученное на основе словаря синонимов, или "по написанию", вычисленное в результате нечеткого сравнения. Сходство слов "по написанию" в работе [13] вычисляется как расстояние Левенштейна между  $n$ -граммами слов. Сравнение строк на основе расстояния Левенштейна (расстояния редактирования) [1, 14] и метрики Джаро-Винклера [14] используется также для выявления нечетких дубликатов в текстах [14].

В данной работе описана векторная модель поиска речевых документов по текстовому запросу. В качестве функции релевантности предложена модифицированная взвешенная косинусная мера, веса которой вычисляются на основе нечеткого сравнения слов распознанного речевого документа и текстового запроса. Нечеткое сравнение слов выполняется посредством нахождения длины наибольшей общей подстроки, либо расстояния Левенштейна. Проведен анализ эффективности поиска речевых документов в зависимости от используемых алгоритмов точного и нечеткого сравнения слов.

### Постановка SDR-задачи

Речевой документ  $d_k$  коллекции  $D$  является набором независимых распознанных слов  $\{w_i^{d_k}\}$ . Текстовый запрос  $Q$  задается набором ключевых слов  $\{q_j\}$ , введенных пользователем.

SDR-задача [1] заключается в вычислении оценки релевантности  $r_k$  документов  $d_k \in D$  запросу  $Q$ :

$$r_k = F(d_k, Q), \quad (1)$$

где  $F$  — функция релевантности  $d_k$  запросу  $Q$ . Требуется сформировать список документов, ранжированный по значению  $r_k$ .

### Векторная модель информационного поиска

Пусть  $T = \{t_i\}$  является множеством слов, встречающихся в коллекции  $D$ . Каждому документу  $d_k \in D$  поставим в соответствие вектор  $\mathbf{d}_k = (v_0, v_1, \dots, v_n)$   $n$ -мерного пространства ( $n = |T|$ ), где  $v_i$  — вес слова  $t_i \in T$  в документе  $d_k \in D$ . Широко используемым

методом взвешивания слов документа является *tf-idf* [1, 15]:

$$v_i = t f_i \cdot i d f_{d_k, i}, \quad (2)$$

где  $t f_i$  — частота вхождения слова  $t_i$  в документ  $d_k$ , а  $i d f_{d_k, i}$  — обратная частота документа, вычисляемая как

$$i d f_{d_k, i} = \log\left(\frac{|D|}{|(d_k \supset t_i)|}\right),$$

где  $|D|$  — число документов в коллекции  $D$ ;  $|(d_k \supset t_i)|$  — число документов, содержащих слово  $t_i$ .

Текстовому запросу  $Q$  поставим в соответствие вектор  $\mathbf{q} = (u_0, u_1, \dots, u_n)$ , определяемый в результате точного сравнения со словами  $t_i \in T$ :

$$u_i = \begin{cases} 0, & t_i \notin Q, \\ 1, & t_i \in Q. \end{cases} \quad (3)$$

Значение функции релевантности  $F(d_k, Q)$  документа  $d_k \in D$  запросу  $Q$  вычисляется как косинусная мера близости между векторами  $\mathbf{d}_k$  и  $\mathbf{q}$  [1]:

$$F(d_k, Q) = \cos(\mathbf{d}_k; \mathbf{q}) = \frac{\sum_{i=1}^n v_i u_i}{\sqrt{\sum_{i=1}^n v_i^2} \cdot \sqrt{\sum_{i=1}^n u_i^2}}. \quad (4)$$

### Взвешенная косинусная мера близости

При поиске по коллекции текстовых документов используется метод точного сравнения строк (2) [1]. Однако указанный метод менее эффективен при поиске по коллекции распознанных речевых документов, так как в тексте присутствуют слова, распознанные с ошибками. Частично учесть при поиске искажения слов, полученные в результате распознавания речевых документов, позволяют методы нечеткого сравнения строк. "Мягкая" косинусная мера [13] имеет вычислительную сложность  $O(|T|^2)$ , что существенно влияет на время обработки запроса.

Введем взвешенную косинусную меру близости:

$$\cos'(\mathbf{d}_k; \mathbf{q}) = \frac{\sum_{i=1}^n s_i v_i u_i}{\sqrt{\sum_{i=1}^n s_i v_i^2} \cdot \sqrt{\sum_{i=1}^n s_i u_i^2}}, \quad (5)$$

где  $s_i$  — вес слова  $t_i \in T$ ,  $u_i = 1, i = 1 \dots n$ .

Для определения веса  $s_i$  необходимо найти меру сходства между словом  $t_i \in T$  и словами запроса  $q_j \in Q$ .

В данной работе для решения рассматриваемой задачи предлагается вычислять значение меры сходства на основе нечеткого сравнения слов. В совре-

менных системах, в которых используют алгоритмы нечеткого сравнения строк, например, для исправления опечаток при вводе запроса пользователем, применяют алгоритмы поиска длины наибольшей общей подстроки и расстояния Левенштейна.

Мера сходства  $g_{t_p, q_j}$  на основе нормированной длины наибольшей общей подстроки равна

$$g_{t_p, q_j} = \frac{|s_{q_p, t_i}|}{\max(|q_j|; |t_i|)}, \quad (6)$$

где  $|s_{q_p, t_i}|$  — длина наибольшей общей подстроки строк  $q_j$  и  $t_i$ , а  $|s|$  — длина строки  $s$ .

Расстояние Левенштейна [1, 14] позволяет вычислять сходство двух строк посредством определения минимального числа операций редактирования, требуемых для преобразования одной строки во вторую. Операциями редактирования являются "вставка", "удаление" и "замена" символа. Для вычисления расстояния Левенштейна применяют алгоритм динамического программирования Вагнера—Фишера [16].

Мера сходства на основе расстояния Левенштейна:

$$g_{t_p, q_j} = 1 - \frac{D(q_j, t_i)}{|q_j| + |t_i|}, \quad (7)$$

где  $D(q_j, t_i)$  — значение расстояния Левенштейна между строками  $q_j$  и  $t_i$ .

Тогда веса  $\{s_i\}$ ,  $i = \overline{1, n}$ , можно вычислить, определив слова  $t_i \in T$ , для которых  $g_{t_p, q_j}$  максимально:

$$s_i = \begin{cases} g_{t_p, q_j}, \exists q_j \in Q: g_{t_p, q_j} = \max_{i=1, \dots, n} (g_{t_p, q_j}) \\ 0, \text{ в противном случае.} \end{cases} \quad (8)$$

Рассмотрим в качестве примера коллекцию из трех документов. В табл. 1 представлено верное содержание документов и содержание, полученное в результате распознавания.

Множество  $T$  всех слов, встречающихся в коллекции, содержит девять слов. В табл. 2 показаны  $tf$ - $idf$  веса (2) слов множества  $T$ ; для вычисления значения  $idf$  использовали десятичный логарифм.

Векторы соответствующих документов равны

$$\mathbf{d}_1 = (0,48; 0,48; 0,48; 0; 0; 0; 0; 0; 0),$$

$$\mathbf{d}_2 = (0; 0; 0; 0,48; 0,18; 0,48; 0; 0; 0),$$

$$\mathbf{d}_3 = (0; 0; 0; 0; 0,18; 0; 0,48; 0,48; 0,48).$$

Пусть требуется вычислить значения релевантности документов относительно запроса "заметали следы". Вычислим значения меры сходства для каждого слова  $t_i \in T$  относительно слов запроса. В табл. 3 представлены значения весов  $s_i$ , которые получены на основе расчета мер сходства (6) и (7).

В соответствии с правилом (8) вектор весовых коэффициентов косинусной меры (5) формируется на основе максимальных значений мер сходства (табл. 3):

$$\mathbf{s}_{\text{sub}} = (0; 0; 0; 0; 0,5; 1; 0; 0; 0),$$

$$\mathbf{s}_{\text{Lev}} = \{0; 0; 0; 0; 0,8; 1; 0; 0; 0\},$$

где  $\mathbf{s}_{\text{sub}}$  — вектор весовых коэффициентов, вычисленных как мера сходства на основе длины наибольшей общей подстроки (6), а  $\mathbf{s}_{\text{Lev}}$  — вектор весовых коэффициентов, вычисленных как мера сходства на основе расстояния Левенштейна (7).

Таблица 1

Содержание речевых документов коллекции

Документ	Верное содержание	Распознанное содержание
Документ 1	Торжественно гарцевавших	Торжественно гонцы ваших
Документ 2	Заметали следы	За мечтали следы
Документ 3	Мечтали о золотом веке	Мечтали по золотому веки

Таблица 2

$tf$ - $idf$  веса слов множества  $T$

Слова	$idf$	Документ 1		Документ 2		Документ 3	
		$tf$	$tf$ - $idf$	$tf$	$tf$ - $idf$	$tf$	$tf$ - $idf$
торжественно	0,48	1	0,48	0	0	0	0
горцы	0,48	1	0,48	0	0	0	0
ваших	0,48	1	0,48	0	0	0	0
за	0,48	0	0	1	0,48	0	0
мечтали	0,18	0	0	1	0,18	1	0,18
следы	0,48	0	0	1	0,48	0	0
по	0,48	0	0	0	0	1	0,48
золотому	0,48	0	0	0	0	1	0,48
веки	0,48	0	0	0	0	1	0,48

Таблица 3

Значения мер сходства

$g_{t_p, q_j}^{substr}$			$g_{t_p, q_j}^{Levenshtein}$		
слова	заметали	следы	слова	заметали	следы
торжественно	0,083	0,083	торжественно	0,5	0,412
горцы	0	0,2	горцы	0,385	0,6
ваших	0,125	0	ваших	0,462	0,5
за	0,25	0	за	0,4	0,286
мечтали	<b>0,5</b>	0,143	мечтали	<b>0,8</b>	0,417
следы	0,125	<b>1</b>	следы	0,462	<b>1</b>
по	0	0	по	0,2	0,286
золотому	0,125	0,125	золотому	0,625	0,462
веки	0,125	0,2	веки	0,5	0,556

**Примечание:**  $g_{t_p, q_j}^{substr}$  — значение меры сходства, вычисленные на основе длины наибольшей общей подстроки (6),  $g_{t_p, q_j}^{Levenshtein}$  — значение меры сходства на основе расстояния Левенштейна (7).

Таблица 4

Значения взвешенного косинуса и ранг документов

Документ	Мера сходства для вычисления $s_i$		Ранг
	Длина наибольшей общей подстроки	Расстояние Левенштейна	
Документ 1	0	0	3
Документ 2	0,936	0,917	1
Документ 3	0,577	0,667	2

Используя (5), получим значения взвешенной косинусной меры (табл. 4).

В результате (табл. 4) речевые документы будут ранжированы по значению релевантности: документ 2, документ 3 и документ 1.

### Эксперимент

В ходе эксперимента оценивали эффективность векторной модели поиска с использованием точного и нечеткого сравнения слов запроса со словами распознанных речевых документов. Анализировали следующие методы.

*Метод 1* — определение релевантности документа запросу на основе точного совпадения слов (3) и косинусной меры (4).

*Метод 2* — определение релевантности документа запросу на основе нечеткого сравнения слов и взвешенной косинусной меры (5), где веса  $s_i$  вычисляются как нормированная длина наибольшей общей подстроки (6).

*Метод 3* — определение релевантности документа запросу на основе нечеткого сравнения слов и взвешенной косинусной меры (5), где веса  $s_i$  определяются на основе расстояния Левенштейна (7).

В различных системах, включающих поиск по текстовым документам, широко распространено использование свободной библиотеки высокоскоростного полнотекстового поиска *Lucene* [17, 18]. Поэтому для сравнения результатов дополнительно выполнялся поиск на основе данной библиотеки.

*Метод 4* — полнотекстовый поиск.

*Метод 5* — полнотекстовый поиск с неточным соответствием слов.

Эффективность поиска оценивали показателями полноты  $R$  и точности  $P$ , полученными в результате эксперимента на разработанной SDR-системе. В соответствии с работой [1] полнота определяется как доля найденных релевантных документов среди всех релевантных, а точность — доля релевантных документов среди найденных.

Оценку указанных показателей выполняли по 253 поисковым запросам на коллекции из 100 речевых документов. Для распознавания речи использована библиотека *pocketsphinx* [19], акустическая и языковая модели [20].

Полученные результаты представлены в табл. 5. Запись "> 0" означает, что релевантными докумен-

тами считались те, для которых оценка релевантности запросу  $r_k$  (1) положительна. Запись " $t = N$ " означает, что релевантными считались документы, имеющие ранг не более  $N$ .

Совместной оценкой полноты и точности является  $F_1$ -мера [1]:

$$F_1 = \frac{2RP}{R+P}.$$

Значения  $F_1$ -меры приведены в табл. 6.

Графики полноты/точности, построенные по полученным значениям, представлены на рисунке. Цифрами обозначены номера исследуемых методов.

Таблица 5

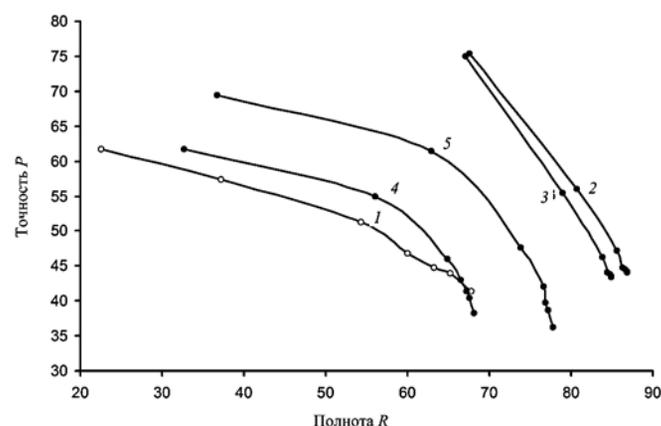
Показатели полноты и точности поиска

Метод	Значения полноты						
	>0	$t = 1$	$t = 2$	$t = 4$	$t = 6$	$t = 8$	$t = 10$
1	67,8	22,5	37,2	54,3	<b>60,0</b>	63,3	65,3
2	86,9	<b>67,5</b>	80,7	85,6	86,2	86,7	86,9
3	84,9	<b>67,1</b>	79,0	83,9	84,5	84,8	84,9
4	68,2	32,7	<b>56,0</b>	64,8	66,5	67,3	67,6
5	77,8	36,8	<b>62,9</b>	73,8	76,6	76,9	77,2
Метод	Значения точности						
	>0	$t = 1$	$t = 2$	$t = 4$	$t = 6$	$t = 8$	$t = 10$
1	41,3	61,7	57,3	51,3	<b>46,8</b>	44,8	43,9
2	44,1	<b>75,4</b>	56,0	47,2	44,7	44,5	44,2
3	43,4	<b>74,9</b>	55,4	46,2	44,1	43,8	43,5
4	38,3	61,7	<b>54,9</b>	46,0	43,0	41,4	40,4
5	36,3	69,4	<b>61,4</b>	47,7	42,1	39,8	38,7

Таблица 6

Значения  $F_1$ -меры ( $F_1$ )

Метод	>0	$t = 1$	$t = 2$	$t = 4$	$t = 6$	$t = 8$	$t = 10$
	$F_1$	$F_1$	$F_1$	$F_1$	$F_1$	$F_1$	$F_1$
1	51,4	33,0	45,1	52,8	<b>52,6</b>	52,5	52,5
2	58,5	<b>71,2</b>	66,1	60,9	58,9	58,8	58,6
3	57,4	<b>70,8</b>	65,1	59,6	57,9	57,8	57,5
4	49,1	42,7	<b>55,5</b>	53,8	52,2	51,2	50,6
5	49,5	48,1	<b>62,2</b>	57,9	54,3	52,4	51,5



Графики полноты/точности исследуемых методов (цифрами обозначены номера методов)

## Заключение

Полученные результаты (табл. 5, 6 и рисунок) позволяют сделать следующие выводы:

1) наилучшее качество поиска речевых документов по текстовому запросу показали методы 2 и 3, включающие нечеткое сравнение слов и вычисленные оценки релевантности посредством взвешенной косинусной меры (5);

2) наихудшее качество поиска из рассмотренных методов показали методы точного сравнения 1 и 4;

3) среднее качество поиска показал метод 5 (полнотекстовый поиск с неточным соответствием слов, реализованный в библиотеке *Lucene*);

4) использование длины наибольшей общей подстроки (6) при вычислении оценки релевантности документа запросу как взвешенной косинусной меры (5) в отличие от расстояния Левенштейна (7) в среднем улучшает полноту поиска на 2 % и точность поиска на 1 %;

5) использование предложенной взвешенной косинусной меры (5) для вычисления оценки релевантности повышает показатель  $F_1$ -меры поиска на 15 % относительно метода полнотекстового поиска с неточным соответствием слов, реализованного в библиотеке *Lucene* [16];

6) как видно из табл. 5 и 6, при использовании векторной модели поиска на основе предложенной взвешенной косинусной меры (5) ложно-определенные системой релевантные документы реже попадают в верхние ранги ранжированного результирующего списка по сравнению с традиционной векторной моделью и полнотекстовым поиском, реализованном в *Lucene*.

### Список литературы

1. Маннинг К. Д., Рагхаван П., Шютце Х. Введение в информационный поиск. М.-СПб.-К.: Вильямс, 2011. 520 с.
2. Лакаев А. С. Контекстные технологии — новое направление развития информационных технологий анализа текстовой информации // Информационные технологии. 2013. № 12. С. 10—16.

3. Larson M., Jones G. J. F. Spoken Content Retrieval: A Survey of Techniques and Technologies // Information Retrieval. 2011. V. 5, N 4—5. P. 235—422.
4. Wechsler M. New Approaches to Spoken Document Retrieval, Information Retrieval // Information Retrieval. 2000. V. 3. P. 173—188.
5. Jones G., Foote J., Jones K. S., Young S. Video mail retrieval using voice: An overview of the stage-2 system // In: van Rijsbergen C., Ed., Proc. of the Final Workshop on Multimedia Information Retrieval (MIRO'95), Electronic Workshops in Computing, Glasgow: Springer, 1995.
6. Wactlar H., Hauptmann A., Witbrock M. Informedia: News-on-demand experiments in speech recognition // In: Proceedings of DARPA Speech Recognition Workshop. Arden House, Harriman, NY. 1996.
7. Wechsler M. Spoken Document Retrieval Based on Phoneme Recognition: PhD Thesis. Diss. No. 12879. ETH Zurich. 1998.
8. Ng K., Zue V. W. Phonetic Recognition for spoken document retrieval // Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing. 1998. V. 1, pp. 325—328.
9. Яшина А. Г. Алгоритм контекстного поиска речевых аудио-файлов на основе фонемного сравнения слов // Advanced Science, 2012. № 1. С. 73—85. URL: [http://www.vyatsu.ru/uploads/file/1210/1\\_\(2\).pdf](http://www.vyatsu.ru/uploads/file/1210/1_(2).pdf) (дата обращения: 12.03.2015).
10. Проzorov Д. Е., Яшина А. Г. Анализ алгоритмов фонемного транскрибирования в задачах контекстного поиска речевых документов // Инфокоммуникационные технологии. 2013. Т. 12, № 4. С. 62—65.
11. Witbrock M., Hauptmann A. G. Speech recognition and information retrieval: Experiments in retrieving spoken documents // In: Proceedings of the DARPA Speech Recognition Workshop. Chantilly Virginia. 1997.
12. Brown M., Foote J., Jones G., Jones K. S., Young S. Open-vocabulary speech indexing for voice and video mail retrieval // In: ACM Multimedia Conference. Boston, MA. 1996.
13. Sidorov G., Gelbukh A., Gomez-Adorno H., Pinto D. Soft similarity and soft cosine measure: similarity of features in vector space model // Computaciony Sistemasy. 2014. V. 18, N 3.
14. Bilenko M., Mooney R., Cohen W., Ravikumar P., Fienberg S. Adaptive name matching in information integration // IEEE Intelligent Systems. 2003. P. 16—23.
15. Salton G., Buckley C. Term-weighting approaches in automatic text retrieval // Information Processing and Management. 1988. V. 24, N 5, pp. 513—523.
16. Wagner R., Fischer M. The String-to-String Correction Problem // Journal of the Association for Computing Machinery. 1974. V. 21, N. 1, pp. 168—173.
17. Hatcher E., Gospodnetic O., McCandless M. Lucene in Action // Manning — 2009.
18. Lucene.NET. URL: <https://www.nuget.org/packages/Lucene.Net> (дата обращения: 12.03.2015).
19. CMU Sphinx. Open Source Toolkit For Speech Recognition // URL: <http://cmusphinx.sourceforge.net> (дата обращения: 12.03.2015).
20. Voxforge-ru-0.2. URL: [http://sourceforge.net/projects/cmusphinx/files/Acoustic and Language Models/Russian Voxforge/](http://sourceforge.net/projects/cmusphinx/files/Acoustic%20and%20Language%20Models/Russian%20Voxforge/) (дата обращения: 12.03.2015).

A. G. Yashina, Assistant, ayashina.pmi@gmail.com, D. E. Prozorov, PhD, Professor, prozorov.de@gmail.com  
Vyatka State University, Kirov

## Spoken Document Retrieval Vector Model Based on Weighted Cosine Measure

*This paper presents a weighted cosine measure as a relevant function in a spoken document retrieval vector model. Measure weights are based on an approximate string matching. A length of the longest common substring and a Levenshtein distance are used in the weighted cosine measure. Search techniques have been evaluated using the collection which is consisted of 100 spoken documents on the russian language. Experiment results are analyzed. Results show that the weighted cosine measure improves a spoken document retrieval based on a vector model in comparison with a standard cosine measure and text retrieval methods of a Lucene search engine library.*

**Keywords:** information retrieval, spoken document retrieval, information retrieval vector model, cosine measure

## References

1. **Manning K. D., Raghavan P., Schütze X.** *Introduction to information search*, M.-SPb.-K.: Williams, 2011, 520 p.
2. **Lakayev A. S.** Contextual technologies — the new direction of development of the text information analysis information technologies, *Information technologies*, 2013, no. 12, pp. 10—16.
3. **Larson M., Jones G. J. F.** Spoken Content Retrieval: A Survey of Techniques and Technologies, *Information Retrieval*, 2011, vol. 5, no. 4—5, pp. 235—422.
4. **Wechsler M.** New Approaches to Spoken Document Retrieval, *Information Retrieval*, 2000, vol. 3, pp. 173—188.
5. **Jones G., Foote J., Jones K. S., Young S.** Video mail retrieval using voice: An overview of the stage-2 system, Ed. van Rijsbergen C., *Proc. of the Final Workshop on Multimedia Information Retrieval (MIRO'95), Electronic Workshops in Computing*, Glasgow, Springer, 1995.
6. **Wactlar H., Hauptmann A., Witbrock M.** Informedia: News-on-demand experiments in speech recognition, *Proc. of DARPA Speech Recognition Workshop*, NY., Arden House, Harriman, 1996.
7. **Wechsler M.** Spoken Document Retrieval Based on Phoneme Recognition: PhD Thesis. Diss. 1998, no. 12879, ETH Zurich.
8. **Ng K., Zue V. W.** Phonetic Recognition for spoken document retrieval, *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 1998, vol. 1, pp. 325—328.
9. **Yashina A. G.** The spoken document retrieval algorithm based on a phonemic word matching, *Advanced Science*, 2012, no. 1, pp. 73—85, available at: [http://www.vyatsu.ru/uploads/file/1210/1\\_\(2\).pdf](http://www.vyatsu.ru/uploads/file/1210/1_(2).pdf)
10. **Prozorov D. E., Yashina A. G.** The analysis of phonemic transcription algorithms in a spoken document retrieval, *Infocommunication technologies*, Samara, 2013, vol. 12, no. 4, pp. 62—65.
11. **Witbrock M., Hauptmann A. G.** Speech recognition and information retrieval: Experiments in retrieving spoken documents, *Proc. of the DARPA Speech Recognition Workshop*, Chantilly Virginia, 1997.
12. **Brown M., Foote J., Jones G., Jones K. S., Young S.** Open-vocabulary speech indexing for voice and video mail retrieval, *ACM Multimedia Conference*, Boston, MA., 1996.
13. **Sidorov G., Gelbukh A., Gomez-Adorno H., Pinto D.** Soft similarity and soft cosine measure: similarity of features in vector space model, *Computaciony Sistemy*, 2014, no. 18 (3).
14. **Bilenko M., Mooney R., Cohen W., Ravikumar P., Fienberg S.** Adaptive name matching in information integration, *IEEE Intelligent Systems*, 2003, pp. 16—23.
15. **Salton G., Buckley C.** Term-weighting approaches in automatic text retrieval, *Information Processing & Management*, 1988, no. 24 (5), pp. 513—523.
16. **Wagner R., Fischer M.** The String-to-String Correction Problem, *Journal of the Association for Computing Machinery*, 1974, vol. 21, no. 1, pp. 168—173.
17. **Hatcher E., Gospodnetic O., McCandless M.** *Lucene in Action*, Manning, 2009.
18. **Lucene.NET**, available at: <https://www.nuget.org/packages/Lucene.Net>
19. **CMU Sphinx.** Open Source Toolkit For Speech Recognition, available at: <http://cmusphinx.sourceforge.net>
20. **Voxforge-ru-0.2**, available at: [http://sourceforge.net/projects/cmusphinx/files/Acoustic and Language Models/Russian Voxforge](http://sourceforge.net/projects/cmusphinx/files/Acoustic%20and%20Language%20Models/Russian/Voxforge)

---

---

### Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: [it@novtex.ru](mailto:it@novtex.ru)

Технический редактор *Е. В. Конова*.

Корректор *Е. В. Комиссарова*.

Сдано в набор 07.07.2015. Подписано в печать 20.08.2015. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ IT915. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансед солюшнз". Отпечатано в ООО "Авансед солюшнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1.

---